目　　录

第一章　差分计算的基本原理

第二章　一维抛物型与双曲型方程的差分方法

第三章 二维、三维流动抛物型与双曲型方程差分研究

第四章 应用计算流体力学专题

第五章 流体力学的有限元素法

第六章 并行计算的基本概念

第七章 并行编程与示范实例

第八章 CFD 并行计算

第一章　差分计算的基本原理

流体力学中的数值计算方法有许多种，如有限差分或有限体积法、有限元法、面元法、特征线法、谱法、积分关系式法、格子法、统计试验法、Monte-Carlo 法、摄动法等。由于流体力学所处理问题的物理模型复杂多样，针对不同的问题可能要采用上述不同的方法。但目前在处理大量工程实际问题中应用最为广泛又较成熟的方法仍是有限差分法（有限体积法）和有限元法。许多方法有其固有的特点，也与其他方法相关。本书主要向大家介绍有限差分法（有限体积法）和有限元法。

从应用上讲，有限差分法较多集中在解决依赖于时间的问题（双曲型方程或抛物型方程），而有限元法则侧重应用于定态问题（椭圆型方程），但这并不是绝对的。现在用空间步进的差分方法解定常问题，以及用有限元法解抛物型问题都已相当普遍。有限元法解双曲型问题也已展开，近年来已取得显著效果，并出现了用有限元法解高超声速粘流问题的报告。

本章主要介绍流体力学基本方程及其数学分类、边界条件的处理方法、有限差分与有限体积法的基本原理等，为以后各章分析差分方法在不同情况下的应用打下基础。

1.1　引言

有限差分法（含有限体积法）是计算流体力学中的主要方法。最早的应用可追溯到1933年，A. Thom 手算了低速绕圆柱流场。有限差分方法简单、方便，易使人误认为只要有大型快速的计算机，有意义的工程问题就可用有限差分法轻易得到解决。其实，差分计算中存在很多问题，如稳定性、收敛性、精度、不规则曲线边界条件处理等以及由方程非线性引起的困难。这说明有限差分方法中还有许多不解之谜，计算方法大有改进的必要；同时也说明了为什么在有限差分法之外还有许多如前面所叙述的其他方法。在有限差分方法被广泛应用的几十年间，差分方法本身也取得了很大的进展，如在复杂边界的网格处理、计算效率的提高、精度的改进、间断面的捕捉等方面都有了长足的进步。

用数值方法解流体力学工程问题时，总是先列出求解问题的控制方程（如 N. S. 方程），然后对方程进行离散，满足具体问题的定解条件，最后在计算机上解出流场。可见，这是经典的理论流体力学、数值计算方法与计算机技术三者的综合。因此，计算流体力学更接近于实验流体力学的实验仿真。只是后者用空间点物理量的测量结果而前者用空间点物理量的计算结果来表示流场。经过近几十年的发展，计算流体力学已接近于实用阶段。目前，三维真实流场的数值模拟和跨声速巡航飞行时飞机全机的气动辅助设计已经实现，飞行器设计、试验及其他相关领域都在发生根本的变革，计算流体力学前景十分宽广。

与风洞实验相比，数值模拟具有以下特点：计算出的流场不受测试器械干扰；影响流场的各物理因素可通过输入量的变化单个加以控制，以分析其影响的程度等，这是风洞中

无法实现的；无粘的假设在风洞中实现不了，但数值计算中可以实现。总之，较之风洞实验，计算流体力学(CFD)有五大优点：

① 研制预研时间大为减少；

② 提供实验无法提供的信息；

③ 更细致和必要的信息；

④ 费用低；

⑤ 能量消耗相对较少。

另一方面，数值模拟本身也面临一些问题，如复杂的物理模型及边界条件常常需要分析实验结果后确定，事先难以列出；网格方面也存在一系列难题。再如，对三维分离流场无粘边界条件效果较好，是通过大量实验比较得来的。对连续介质流动问题，由于差分离散，引进人工粘性，影响了解的准确性。对于湍流工程问题，限于计算机的水平，目前尚不可能完全直接模拟，只能借助于工程模型。控制方程的非线性也带来计算上的问题。由此可见，为了使计算流体力学更好地为工程服务，数值方法与计算机均需进一步的发展；严格地用定解条件解偏微分方程，首先要对问题进行适定性论证，而三维真实流场的严格论证目前尚未见到，且非线性方程(如拟线性的 Euler 或 N. S. 方程)严密的数学处理也存在着困难，所以大量工程实际问题的数值计算仍属于“后验”的范畴，计算所得的结果应得到实验的验证。

下面从流体动力学的控制方程及其数学分类出发来介绍计算流体力学的基本原理。

1.2　流体力学控制方程组

典型的连续介质流体力学控制方程是 Navier-Stokes 方程组(简称 N. S. 方程组)。它包括质量守恒、动量守恒和能量守恒方程，对可压流还包括状态方程。如果是层流流动，方程组是封闭的。对湍流流动通常要补充湍流的特征方程或工程模型。为了使讨论的问题不太复杂，又不失一般性，我们先对二维直角坐标系下的方程进行分析。三维数值计算方法不难从二维加以延伸。

1.2.1　积分形式与微分形式的 N. S. 方程组

设 p,ρ,T 分别为压力、密度、温度；u,v 为笛卡尔坐标系(x,y)的两个速度分量；E_1，μ_l,μ_t,k 分别为单位体积的总能量，粘性系数，湍值粘性系数和传热系数；γ,P_r,P_{rt}分别为比热比、普朗特数、湍流普朗特数；$u_*,M_*,\rho_*,T_*,\mathrm{Re}_*,\mu_*,L_*$分别为参考速度、马赫数、密度、温度、雷诺数、粘性系数和长度。$M_*=u_*/\sqrt{\gamma RT_*}$，$\mathrm{Re}_*=(\rho_* u_* L_*)/\mu_*$；$\mu_*$为一个大气压下，0℃时空气的粘性系数。定义如下无量纲变量：

$$\bar{x}=\frac{x}{L_*},\bar{y}=\frac{y}{L_*},\bar{u}=\frac{u}{u_*},\bar{v}=\frac{v}{u_*},\bar{\rho}=\frac{\rho}{\rho_*},\bar{T}=\frac{T}{T_*},\bar{p}=\frac{p}{\rho_* u_*^2},$$

$$\bar{t}(\text{时间})=\frac{t}{(L_*/u_*)},\bar{E}_1=\frac{E_1}{\rho_* u_*^2},\bar{\mu}_l=\frac{\mu_l}{\mu_*},\bar{\mu}_t=\frac{\mu_t}{\mu_*}。$$

假设无源，无彻体力，无量纲物理量去掉上划线，则非定常条件下完全气体的无量纲

守恒形式的层流 N. S. 微分方程组为：

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \tag{1.1}$$

其中：$U=\begin{pmatrix}\rho\\ \rho u\\ \rho v\\ E_1\end{pmatrix}$，$F=F_I-F_V$，$G=G_I-G_V$，下标"$I$"表示无粘部分，"$V$" 表示粘性部分。

$$F_I=\begin{pmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ (E_1+p)u\end{pmatrix},G_I=\begin{pmatrix}\rho v\\ \rho uv\\ \rho v^2+p\\ (E_1+p)v\end{pmatrix},$$

$$F_V=\frac{1}{\mathrm{Re}_\infty}\begin{pmatrix}0\\ \tau_{xx}\\ \tau_{xy}\\ \tau_{xx}u+\tau_{xy}v-q_x\end{pmatrix},G_V=\frac{1}{\mathrm{Re}_\infty}\begin{pmatrix}0\\ \tau_{yx}\\ \tau_{yy}\\ \tau_{yx}u+\tau_{yy}v-q_y\end{pmatrix},$$

$$E_1=\frac{p}{\gamma-1}+\frac{1}{2}\rho(u^2+v^2),\ \tau_{xx}=(2\mu+\lambda)u_x+\lambda v_y,$$

$$\tau_{yy}=(2\mu+\lambda)v_y+\lambda u_x,\ \tau_{xy}=\tau_{yx}=\mu(u_y+v_x),$$

$$q_x=-\frac{\mu}{P_r}\cdot\frac{C_p}{(\gamma-1)M_\infty^2}\cdot T_x=-\frac{k}{(\gamma-1)M_\infty^2}\cdot T_x$$

$$q_y=-\frac{\mu}{P_r}\cdot\frac{C_p}{(\gamma-1)M_\infty^2}\cdot T_y=-\frac{k}{(\gamma-1)M_\infty^2}\cdot T_y$$

λ 与 μ 通常满足 Stokes 关系式：$3\lambda+2\mu=0$，即 $\lambda=-\frac{2}{3}\mu$。

应用 Gauss 散度定律，(1.1)可写成积分形式：

$$\frac{\partial}{\partial t}\int_D U\mathrm{d}S+\int_\Gamma \boldsymbol{H}\cdot\boldsymbol{n}\mathrm{d}l=0 \tag{1.2}$$

式中：Γ 是固定域 D 的边界；$\boldsymbol{H}$ 是相应的矢通量，$\boldsymbol{n}$ 是单位外法向矢量。

对层流流动，粘性系数 μ 随 T 而变化，这可用苏士兰公式(Sutherland)表示：

$$\frac{\mu}{\mu_0}=\left(\frac{T}{T_0}\right)^{\frac{3}{2}}\cdot\frac{T_0+T_S}{T+T_S} \tag{1.3}$$

式中 $T_0=273.16\mathrm{K}$；T_S 为 Sutherland 常数，对空气取 $T_S=110.4\mathrm{K}$；μ_0 为 1atm，273.16K 下空气的粘性系数。

对于湍流流动，工程上采用时间平均值法，得雷诺平均 N. S. 方程。若忽略密度脉动及其他小量脉动的影响，按布辛涅斯克(Boussinesq)假设，雷诺应力项中：

$$-\overline{\rho u'v'}=\mu_t\frac{\partial u}{\partial y},\ -\overline{\rho u'^2}=-2\mu_t\frac{\partial v}{\partial y}+\frac{2}{3}\rho K$$

式中 $K=\frac{1}{2}(\overline{u'^2}+\overline{v'^2})$是湍流动能。在上述假设下仍可使用上述(1.1)或(1.2)求解，但方程中的 μ 为：

$$\mu = \mu_t + \mu_l \tag{1.4}$$

通常 $\mu_t \gg \mu_l$。μ_t 要用工程湍流模型来求得。湍流模型的合理选取与确定将会对计算结果产生一定的影响，但就数值计算而言，不同的湍流模型一般不会在计算中造成严重困难，因此湍流模型的分析、选取不是本书要解决的问题。

对于复杂的几何边界问题，常采用贴体坐标系统。

设一般曲线坐标系为(ξ,η)，引入如下坐标变换：

$$\begin{cases} \tau = t \\ \xi = \xi(t,x,y) \\ \eta = \eta(t,x,y) \end{cases}$$

则通过坐标变换，方程(1.1)变为：

$$\frac{\partial \overline{U}}{\partial \tau} + \frac{\partial \overline{F}}{\partial \xi} + \frac{\partial \overline{G}}{\partial \eta} = 0 \tag{1.5}$$

其中：

$$\overline{U} = J^{-1}U \tag{1.6}$$

$$\left.\begin{aligned} \overline{F} &= J^{-1}[\xi_t U + \xi_x F + \xi_y G] \\ \overline{G} &= J^{-1}[\eta_t U + \eta_x F + \eta_y G] \end{aligned}\right\} \tag{1.7}$$

J 为坐标变换的雅可比行列式，即

$$J = \left|\frac{\partial(\xi,\eta)}{\partial(x,y)}\right| = \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} \tag{1.8}$$

比较(1.1)和(1.5)可见，一般坐标系(ξ,η,ζ)和笛卡尔坐标系(x,y,z)下的 N.S. 方程的形式是类同的。

方程(1.1)，(1.2)和(1.5)将是我们分析问题的控制方程。尤其是(1.1)，以后经常要引用。它们代表了流体动力学控制方程的基本特性。

1.2.2 N.S. 方程组的几种近似简化形式

用 N.S. 方程组描述的实际工程问题，通常属于三维复杂几何边界问题，求解较为困难，因而在某些条件下对其进行近似简化是必要的。这些简化处理可以在某些问题中全场应用，也可以在另一些问题中应用于某些局部区域。对后者而言，整个流场不同的区域，按当地条件采用不同的简化处理，然后将不同区段衔接起来，得到全流场的解，使整个流场的计算效率提高。

一、无粘流动近似

1. 欧拉(Euler)方程组

令(1.1)中的 $Re \to \infty$，则得如下欧拉方程组(包括无粘的质量、动量与能量守恒方程)：

$$\frac{\partial U}{\partial t} + \frac{\partial F_I}{\partial x} + \frac{\partial G_I}{\partial y} = 0 \tag{1.9}$$

2. 无粘匀熵流(无旋流)

有速度势存在，可化为全速势方程。定常全速势方程为：

$$(a^2 - \phi_x^2)\phi_{xx} - 2\phi_x\phi_y\phi_{xy} + (a^2 - \phi_y^2)\phi_{yy} = 0 \tag{1.10}$$

式中 ϕ 为速度势，a 是当地声速。

3. 小扰动速势方程

定常小扰动亚声速与超声速速势方程为：

$$(1 - M_\infty^2)\phi'_{xx} + \phi'_{yy} = 0 \tag{1.11}$$

跨声速小扰动速势方程：

$$(1 - M_\infty^2 - \frac{\gamma + 1}{U_\infty}M_\infty^2\phi'_x)\phi'_{xx} + \phi'_{yy} = 0 \tag{1.12}$$

式中 ϕ' 是小扰动速度势。

二、粘性流简化近似

1. 边界层近似

在定常 N.S. 方程中略去 $1/\mathrm{Re}^{\frac{1}{2}}$ 及更小量级的项而得边界层方程组（无因次形式）：

$$\left.\begin{aligned}
&\frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial y}(\rho v) = 0 \\
&\rho\left(u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y}\right) = -\frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\partial}{\partial y}\left(\frac{\mu}{\mathrm{Re}_\infty}\frac{\partial u}{\partial y}\right) \\
&\rho\left(u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y}\right) = (\gamma - 1)M_\infty^2 u\frac{\mathrm{d}p}{\mathrm{d}x} + \frac{(\gamma - 1)M_\infty^2\mu}{\mathrm{Re}_\infty}\left(\frac{\partial u}{\partial y}\right)^2 + \frac{\partial}{\partial y}\left(\frac{\mu}{\mathrm{Re}_\infty \mathrm{Pr}}\frac{\partial h}{\partial y}\right) \\
&\mathrm{d}h = C_p\mathrm{d}T,\ p = (\gamma - 1)\rho e
\end{aligned}\right\} \tag{1.13}$$

h 是焓，C_p 是定压比热。边界层方程的求解已有成熟的方法。

2. 薄层 N.S. 方程

在非定常 N.S. 方程组中，将粘性耗散项中沿主流方向导数的一部分或全部略去而得到薄层 N.S. 方程（TLNS）。通常设 x 向（ξ 向）为主流方向。

由（1.1）得：

$$\frac{\partial U}{\partial t} + \frac{\partial F_I}{\partial x} + \frac{\partial G_I}{\partial y} = \frac{\partial G'_V}{\partial y} \tag{1.14}$$

由（1.5）得：

$$\frac{\partial \bar{U}}{\partial t} + \frac{\partial \bar{F}}{\partial \xi} + \frac{\partial \bar{G}}{\partial \eta} = \frac{\partial \bar{G}'_V}{\partial \eta} \tag{1.15}$$

3. 抛物化 N.S. 方程组（PNS）

定常情况下的薄层 N.S. 方程组即为抛物化 N.S. 方程组。

由（1.14）得：

$$\frac{\partial F_I^*}{\partial x} + \frac{\partial G_I}{\partial y} = \frac{\partial G_V^*}{\partial y} \tag{1.16}$$

由（1.15）得：

$$\frac{\partial \bar{F}_I^*}{\partial \xi} + \frac{\partial \bar{G}_I}{\partial \eta} = \frac{\partial \bar{G}_V^*}{\partial \eta} \tag{1.17}$$

上述方程中 G'_V 指的是 G_V 中，凡含 x 或 ξ 向导数的项可以略去，或适当保留。但在 G'_V

中所有的 x 或 ξ 向导数均略去。

TLNS 与 PNS 的差别在于：首先，PNS 无时间相关项；其次，粘性项中略去的项也有不同，TLNS 中的粘性项里有时也包括对流场影响很大的 x 或 ξ 向的某些导数。

必须指出，上述 PNS 近似，要求流场中粘性层以外的流动是超声速的，而且来流速度方向恒为正，并且这个方向正是步进的方向。为保证方程求解适定，在粘性亚声速薄层方程(1.14)中 F_l 中的压强 p 要作适当的处理，因此在(1.16)与(1.17)中均注上了 * 号。对某些不出现较大分离的超声速粘性流场，用抛物化 N. S. 方程空间步进求解，可使计算效率大为提高。

此外还有粘性激波层近似，主要用于计算高超声速钝头体问题。应当说明，在对N. S. 方程进行简化时哪些项(尤其是 G_V 中)该略去，哪些该保留，通常按问题的性质进行取舍，带有一定的灵活性。有人指出，如果严格按量级概念进行取舍，得到的简化方程常常与一些资料中介绍的不完全一致。所以 N. S. 方程的简化形式有多种，上面只介绍了一些常用的形式。

三、流体力学控制方程组的使用范围

上述流体力学控制方程的各个形式均是在连续介质的前提下导出的。对于自由分子流范畴的问题，诸如卫星阻力、强激波内部结构、绕小物体的低密度流等问题，不能用它们作为控制方程组。

对于可压流动的外流问题，如飞机、导弹的绕流，如果流场是接触流，或只有弱分离流动区域，那么用解 Euler 方程的方法，或用 Euler 方程与边界层方程耦合的方法求解比较合适，花费较 N. S. 方程要少，精度达到工程要求。对一些问题，尤其是跨声速流动问题，目前甚至采用全位势方程或小扰动位势方程求解，使问题能快速、低耗地得到解决。目前用解 Euler 方程得出全机或全弹正常飞行下表面压强、马赫数分布，或用位势方程解出全机跨声速飞行时的压强与马赫数分布已大量出现。事实证明，解位势及 Euler 方程仍是工程上可采用的方法。随着计算机性能的提高与计算方法的改进，用更复杂的方程求解工程问题将日益增多。

通常在下列情况下才必须采用雷诺平均的 N. S. 方程组(湍流流动)或 N. S. 方程组来求解：

1. *低雷诺数流动*

流场中粘性耗散起主要作用。这包括雷诺数为 100 量级，马赫数为 1 的量级。如绕热线测速计热线的流动及钝头体超声速流场中，脱体弓形激波与物体间的粘性流动等，此时粘性及热效应起主要作用；在这样的流动中，对 N. S. 方程的简化会引起较大的误差。

2. *中等或大雷诺数流动*

这种流动中，无粘流区与粘性流区有较强的相互干扰，包括较强的分离流动；或者是湍流流动，包括后体尾迹流、底部流、压缩拐角流、火箭发动机燃烧室内流、大曲率的弯曲管道流、激波边界层干扰流等。应当说，湍流与分离流是其两大特征，此时用 N. S. 方程组求解是必须的。

即使对上述两种情况下的流场，依据流动特点对全流场或某些流区采用不同程度的简化，也是可行的。针对求解问题的性质，选用正确有效的控制方程组是首先要注意的。

四、复杂流场的控制方程

对一些更复杂的流动，如多相流、加质量燃烧流等等，控制方程显然比已介绍的N.S.方程组更复杂。以质量守恒和动量守恒的微分方程为例，多相流质量守恒方程为：

$$\frac{\partial}{\partial t}(\rho_i \alpha_i) + \frac{\partial}{\partial x}(\rho_i \alpha_i u_i) + \frac{\partial}{\partial y}(\rho_i \alpha_i v_i) = M_i \tag{1.18}$$

式中"i"是指多相流中某一相的组分，α_i 是体积百分比，M_i 是源项。显然，流动中有多少组分，上述质量方程即可写成多少个。动量守恒方程为：

$$\frac{\partial}{\partial t}(\rho_i \alpha_i u_{il}) + \mathrm{div}(\rho_i \alpha_i u_{il} v_i) = \alpha_i(-I \cdot \mathrm{grad}(p) + B_{il}) + F_{il} + I_{il} \tag{1.19}$$

式中"l"指某一个坐标方向，"B_{il}"是质量力分量，F_{il}是粘性力分量，I_{il}是干扰力分量。除了不同的坐标方向应列出对应的方程外，对不同的组分也应列出对应的方程。这种方程当然比已列出的 N.S. 方程更复杂。

可见，经典的 N.S. 方程组不可能完全控制流体力学中的全部流动。我们选择二维 N.S. 方程组(1.1)等作为典型控制方程，其目的是它的数学性质具有代表性。通过对 N.S. 方程组的分析、论证，基本上可以包括流体力学中大部分有意义的流动的数值求解方法。如上述方程(1.18)、(1.19)，虽然比(1.1)的方程复杂些，但方程的数学性质与处理方法没有本质的变化。因此，选用二维直角坐标下的 N.S. 方程组及其简化形式作为控制方程的代表进行研究是可行的，有意义的。

我们在(1.1)或(1.5)式所给出的完全气体 N.S. 方程组是强守恒形式，关于这一点将在第二章再作说明。

1.3 流体力学控制方程组的数学性质与求解

流体力学控制方程是拟线性偏微分方程，其时间导数项是线性的，而空间导数项常常是非线性的，因此计算流体力学或流体流动中的数值方法，是偏微分方程数值解法的一部分。为此，首先要对控制方程的数学性质有一个较好的了解。

1.3.1 偏微分方程的分类及其意义

一、分类的意义

偏微分方程一般划分为双曲型、抛物型、椭圆型三种类型。不同类型的方程所描述的流场的主要特征与物理背景很不一样，求解时定解条件的提法也不同，数值计算的处理方法也大相径庭。因此，本书首先按数理方程的基本原理对控制方程组在不同条件下是什么类型进行分类。三种类型方程的基本差别如下：

1. 三种类型微分方程解的适定性要求对定解条件有不同的提法

以椭圆型方程为例：

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y),(x,y) \in D \tag{1.20}$$

提第一类边值问题：

$$u(x,y) = g(x,y),\ (x,y \in \partial D) \tag{1.21}$$

其中 D 是边界为 ∂D 的有界区域。这样的问题是适定的，即解存在、唯一，且解稳定即连续依赖于定解条件或其他附属数据。但如果对(1.20)提柯西(Cauchy)问题，即纯初值问题，通常是不适定的。哈特玛(Hadamard)指出，此时只要初始条件稍有变化，解就会出现大的变动，即不适定。在数值求解过程中，定解条件表达中误差的出现是不可避免的，较小的误差若引起解的很大变动，将会使解失去意义。另外，对双曲型方程如只提边值问题，也是不适定的。总之，三种不同类型的方程，其定解条件有相应的正确提法，进而对数值解的定解条件也有相应的规则。若违背这些规则，就得不到正确的解。

2. 三类方程解的光滑性不同，对定解条件的光滑性的要求也不一样

椭圆型、抛物型方程的解是充分光滑的，因此对定解条件的光滑性要求不高。双曲型方程允许有所谓弱解存在(如流场中的激波)，即解的一阶导数可以不连续，因而对定解条件的光滑性要求甚高。这也正是有限元法解双曲型方程困难较多的原因之一。

3. 三类方程的影响区和依赖区不一样

双曲型和抛物型方程控制的流场，某一点的影响区是有界的，可以步进求解。如双曲型方程求解时，为了与影响区的特征一致，采用迎风差分格式比较适宜。而椭圆型方程的影响范围遍及全场，必须全场计算，所选用的差分方法也要相应采用中心差分格式。

从上面的叙述可知，应当对控制微分方程进行数学分类，然后根据其不同的属性分别探讨求解的数值方法。

二、偏微分方程的分类方法

这里不准备进行严格的数学推演，彼德罗夫斯基所著《偏微分方程讲义》一书有这方面较细致的叙述。这里只介绍与我们有关的二阶线性偏微分方程及一阶偏微分方程组的分类方法。这些内容基本上可用于流体力学控制方程的分类。

1. 二阶线性微分方程的分类

考虑方程

$$a_{11}\frac{\partial^2 u}{\partial x^2}+2a_{12}\frac{\partial^2 u}{\partial x\partial y}+a_{22}\frac{\partial^2 u}{\partial y^2}+b_1\frac{\partial u}{\partial x}+b_2\frac{\partial u}{\partial y}+cu=f \tag{1.22}$$

式中 $a_{11},a_{12},a_{22},b_1,b_2,c$ 均不是 $u,\frac{\partial u}{\partial x},\frac{\partial^2 u}{\partial x^2},\frac{\partial u}{\partial y},\frac{\partial^2 u}{\partial y^2}$等的函数。

设有判别式：

$$\Delta=a_{12}^2-a_{11}a_{22} \tag{1.23}$$

当 $\Delta>0$ 时，(1.22)是双曲线方程；$\Delta=0$ 时，(1.22)是抛物型方程；$\Delta<0$ 时，(1.22)是椭圆型方程。

通过上述分析可见，方程的类型由方程中最高阶导数项系数的性质决定，方程中的低阶导数项系数一般不影响方程的数学类型。如果按特征线来分析，$\Delta>0$ 时有两条特征线(特征根为相异实根)，即是双曲型方程；抛物型方程只有一条相重合的特征线；椭圆方程因为特征根为虚数，不存在特征线。因为椭圆型方程不具有类时间的特性，因此必须与双曲、抛物线方程区别开来。以波动方程为例：

$$\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}=H(x,y) \tag{1.24}$$

$a_{11}=1,a_{12}=0,a_{22}=-1,\Delta=1>0$，故(1.24)是双曲型方程。

2. 偏微分方程组的数学分类

考虑一阶偏微分方程组

$$A\frac{\partial U}{\partial x}+B\frac{\partial U}{\partial y}=H \tag{1.25}$$

式中$U=(u_1,u_2,\cdots,u_n)^T$和$H=(H_1,H_2,\cdots,H_n)^T$为n阶列矢量。A和B为$n\times n$矩阵。(1.25)的特征方程为：

$$\det(\sigma_{1i}a_{ij}+\sigma_{2i}b_{ij})=0 \tag{1.26}$$

a_{ij},b_{ij}是A,B矩阵中各对应元素。由(1.26)可解得特征根为

$$\lambda_i=-\frac{\sigma_{1i}}{\sigma_{2i}},\ i=1,2,\cdots,n$$

对于双曲型方程λ_i为n个相异实根；对于椭圆型方程λ_i为虚根或复数根。

应当指出，实际上控制方程组是比较复杂的。若在n个特征根中有若干相异实根，若干重根，则是双曲一抛物型方程组，方程组整体上属于抛物线的。若在n个特征根中有若干复数根，若干重根，则是抛物一椭圆型方程组，方程组整体上属于椭圆型的。另外，定常情况下的N.S.方程组是二阶偏导数的联立方程，也可以将其化成若干个一阶方程组联立，其特征由最高阶偏导数所组成的一次齐式的主部所决定，但同时又要考虑次特征的影响。所谓次特征，即略去最高阶导数后，低阶导数所组成的方程组的特征；对N.S.方程组就是指略去了二阶导数的Euler方程组的特征。

1.3.2 流体力学控制方程组的分类

对N.S.方程组及其简化形式进行严格的数学分类有时是较为困难的。我们通过下列假设对方程进行分类。大量的数值计算表明，这样分类是可行的、合理的。

(1) 方程组的各个方程允许有不同的类型，且在求解过程中各自保持自己的类型，从而要求我们在提定解条件时要谨慎，使之与对应的方程的类型相适应。

(2) 流场的数学特征应由控制方程组中占主导地位的几个方程所决定，流场中不同区域方程的特征可能会变化。如一定攻角下气流绕薄翼流动的全流场计算，在边界层以外流场具有双曲型的特征，而边界层内流场具有抛物型特征。

前面已说明，由两个方程组成的方程组中，若有一个方程是抛物型，另一个是双曲型的，那么整个方程组即是抛物型的。现举例说明。

$$u_x+v_y=0 \tag{1.27}$$

$$v_x+u_y=bv_{yy} \tag{1.28}$$

显然，(1.27)是双曲型方程，(1.28)是抛物型方程。将二者写成矢量形式：

$$\begin{pmatrix}u\\v\end{pmatrix}_x+\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix}_y=\begin{pmatrix}0\\bv_{yy}\end{pmatrix}$$

由(1.27)+(1.28)得

$$(v+u)_x+(u+v)_y=bv_{yy} \tag{1.29}$$

由(1.28)−(1.27)得

$$(v-u)_x+(u-v)_y=bv_{yy} \tag{1.30}$$

将(1.29)和(1.30)写成向量形式得：

$$\begin{pmatrix} u+v \\ v-u \end{pmatrix}_x + \begin{pmatrix} u+v \\ u-v \end{pmatrix}_y = b\begin{pmatrix} v_{yy} \\ v_{yy} \end{pmatrix}$$

令 $r=u+v, s=u-v$，则上式变为：

$$\begin{pmatrix} r \\ -s \end{pmatrix}_x + \begin{pmatrix} r \\ s \end{pmatrix}_y = \frac{b}{2}\begin{pmatrix} r-s \\ r-s \end{pmatrix}_{yy} \tag{1.31}$$

显然，(1.31)中的两个方程都是抛物型的。

(3) 若方程组中各方程由多个变量所组成，应由方程的物理特性确定一个主变量，并由此对方程进行分类。例如定常质量方程 $u\frac{\partial\rho}{\partial x}+\rho\frac{\partial u}{\partial x}+v\frac{\partial\rho}{\partial y}+\rho\frac{\partial v}{\partial y}=0$ 中反映了质量守恒，显然此方程中 ρ 是主要变量。其他如流动方向的动量守恒方程中，流速 u 应是主要变量。方程中各导数项前的各个变量，采用"冻结"的办法，取当地值进行计算与分析。

下面按照上述假设先对一些简化 N.S. 方程组分类。

1. 定常可压流 N.S. 方程组

由(1.1)得：

$$\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0 \tag{1.32}$$

这是由含有 $\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial y^2}, \cdots$ 等二阶偏导数的四个方程组成的方程组。为方便起见，将(1.32)写成无因次的非守恒形式：

$$\left.\begin{aligned}
& u\frac{\partial\rho}{\partial x} + v\frac{\partial\rho}{\partial y} + \rho\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right) = 0 \\
& u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{1}{\rho}\frac{\partial p}{\partial x} = \\
& \qquad \frac{1}{\rho \mathrm{Re}_\infty}\left\{\frac{\partial}{\partial x}\left[\mu\left(\frac{4}{3}\frac{\partial u}{\partial x} - \frac{2}{3}\frac{\partial v}{\partial y}\right)\right] + \frac{\partial}{\partial y}\left(\mu\frac{\partial v}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mu\frac{\partial u}{\partial y}\right)\right\} \\
& u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{1}{\rho}\frac{\partial p}{\partial y} = \\
& \qquad \frac{1}{\rho \mathrm{Re}_\infty}\left\{\frac{\partial}{\partial y}\left[\mu\left(\frac{4}{3}\frac{\partial v}{\partial y} - \frac{2}{3}\frac{\partial u}{\partial x}\right)\right] + \frac{\partial}{\partial x}\left(\mu\frac{\partial v}{\partial x} + \mu\frac{\partial u}{\partial y}\right)\right\} \\
& \rho\left(u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y}\right) - \left(u\frac{\partial p}{\partial x} + v\frac{\partial p}{\partial y}\right)(\gamma-1)M_\infty^2 = \\
& \qquad \frac{\Phi(\gamma-1)M_\infty^2}{\mathrm{Re}_\infty} + \frac{1}{\mathrm{Re}_\infty}\left[\frac{\partial}{\partial x}\left(\frac{\mu}{\mathrm{Pr}}\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\mu}{\mathrm{Pr}}\frac{\partial T}{\partial y}\right)\right] \\
& \Phi = \mu\left[2\left(\frac{\partial u}{\partial x}\right)^2 + 2\left(\frac{\partial v}{\partial y}\right)^2 + \right. \\
& \qquad \left. 2\frac{\partial v}{\partial x}\frac{\partial u}{\partial y} - \frac{2}{3}\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2\right]
\end{aligned}\right\} \tag{1.33}$$

对(1.33)进行一些变换，令：

$$\frac{\partial u}{\partial x}=x_u,\frac{\partial u}{\partial y}=y_u,\frac{\partial v}{\partial x}=x_v,\frac{\partial v}{\partial y}=y_v,$$

$$\frac{\partial T}{\partial x}=x_T,\frac{\partial T}{\partial y}=y_T,(\gamma-1)M_\infty^2=N$$

并用状态方程 $P=\rho T$ 消去(1.33)中的压强项，(1.33)转化为：

$$A\frac{\partial z}{\partial x}+B\frac{\partial z}{\partial y}=H \tag{1.34}$$

其中 $z=(\rho,u,x_u,y_u,v,x_v,y_v,T,x_T,y_T)^T$

$$A=\begin{bmatrix} u & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ -\frac{T}{\rho} & 0 & \frac{4\mu}{3\rho\mathrm{Re}_\infty} & 0 & 0 & 0 & -\frac{2\mu}{3\rho\mathrm{Re}_\infty} & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{\mu}{\rho\mathrm{Re}_\infty} & 0 & \frac{\mu}{\rho\mathrm{Re}_\infty} & 0 & 0 & 0 & 0\\ uT & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\mu}{\rho\mathrm{Re}_\infty P_r} & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B=\begin{bmatrix} v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{\mu}{\rho\mathrm{Re}_\infty} & 0 & \frac{\mu}{\rho\mathrm{Re}_\infty} & 0 & 0 & 0 & 0\\ -\frac{T}{\rho} & 0 & \frac{2\mu}{3\rho\mathrm{Re}_\infty} & 0 & 0 & 0 & \frac{4\mu}{3\rho\mathrm{Re}_\infty} & 0 & 0 & 0\\ vT & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\mu}{\rho\mathrm{Re}_\infty P_r}\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$
H = \begin{pmatrix} -\rho(x_u + y_v) \\ ux_u + vy_u + x_T \\ ux_v + vy_v + y_T \\ \rho(ux_T + vy_T) - \rho N(ux_T + vy_T) - \dfrac{N}{\mathrm{Re}_*}\Phi \\ x_u \\ x_v \\ x_T \\ y_u \\ y_v \\ y_T \end{pmatrix}
$$

(1.34)的特征方程是

$$
\det(\sigma_{1i}a_{ij} + \sigma_{2i}b_{ij}) = 0 \tag{1.35}
$$

由此求得特征根 $\lambda_i = -\left(\dfrac{\sigma_{1i}}{\sigma_{2i}}\right)$:

$$
\lambda_{1,2,3} = 0,\ \lambda_4 = \frac{v}{u},\ \lambda_{5,6,7} = i,\ \lambda_{8,9,10} = -i \tag{1.36}
$$

由此可以断定定常可压流 N. S. 方程组是椭圆型的，但又具有双曲型的性质。这种以椭圆型为主要特征又伴有部分双曲型性质的方程组，方程类型不明确，不便于求解。事实上，每一方程的性质主要由其最高阶导数项来确定，因此可用冻结系数方法及(1.23)式认定动量方程与能量方程皆是椭圆型的。

2. 定常可压流 Euler 方程组

方程形式

$$
\frac{\partial F_I}{\partial x} + \frac{\partial G_I}{\partial y} = 0 \tag{1.37}
$$

改写为以下非守恒形式：

$$
\left.\begin{aligned}
& u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + \rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0 \\
& u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0 \\
& u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{1}{\rho}\frac{\partial p}{\partial y} = 0 \\
& u\frac{\partial p}{\partial x} + v\frac{\partial p}{\partial y} + a^2\rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0
\end{aligned}\right\} \tag{1.38}
$$

改写成：

$$
A'\frac{\partial U}{\partial x} + B'\frac{\partial U}{\partial y} = 0 \tag{1.39}
$$

其中，$U = (\rho, u, v, p)^T$，

$$A' = \begin{pmatrix} u & \rho & 0 & 0 \\ 0 & u & 0 & \frac{1}{\rho} \\ 0 & 0 & u & 0 \\ 0 & a^2\rho & 0 & u \end{pmatrix}, \quad B' = \begin{pmatrix} v & 0 & \rho & 0 \\ 0 & v & 0 & 0 \\ 0 & 0 & v & \frac{1}{\rho} \\ 0 & 0 & a^2\rho & v \end{pmatrix}$$

(1.39)的特征方程为 $\det(\sigma_1 a'_{ij} + \sigma_2 b'_{ij}) = 0$，即：

$$\det\left(\frac{\sigma_1}{\sigma_2} a'_{ij} + b'_{ij}\right) = 0$$

$$\det(\lambda a'_{ij} - b'_{ij}) = 0$$

$$\begin{vmatrix} u\lambda - v & \rho\lambda & -\rho & 0 \\ 0 & u\lambda - v & 0 & \frac{\lambda}{\rho} \\ 0 & 0 & u\lambda - v & -\frac{1}{\rho} \\ 0 & a^2\rho\lambda & -a^2\rho & u\lambda - v \end{vmatrix} = 0$$

$$(u\lambda - v)^2[(u\lambda - v)^2 - a^2\lambda^2 - a^2] = 0$$

由此解得：

$$\lambda_{1,2} = \frac{v}{u}, \lambda_{3,4} = \frac{uv \pm \sqrt{(u^2 + v^2) - a^2}}{u^2 - a^2} \tag{1.39a}$$

因此，在超声速流动中，$(u^2+v^2) > a^2$，定常 Euler 方程是双曲抛物型方程；声速流则是抛物型方程；亚声速流则是椭圆—抛物型方程。亚声速流 Euler 方程组分类数学性质不明确，具有椭圆型性质。

3. 定常抛物化 N.S. 方程组

以(1.16)式为例。它是(1.33)式中，以流动方向 x 为主流方向，将方程中粘性耗散项中的含 x 向导数项略去而得(假设 μ 的变化不影响方程性质)：

$$\left.\begin{aligned} & u\frac{\partial\rho}{\partial x} + v\frac{\partial\rho}{\partial y} + \rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0 \\ & u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{1}{\rho}\frac{\partial p}{\partial x} = \frac{\mu}{\rho \mathrm{Re}_*}\frac{\partial^2 u}{\partial y^2} \\ & u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{1}{\rho}\frac{\partial p}{\partial y} = \frac{4\mu}{3\rho \mathrm{Re}_*}\frac{\partial^2 v}{\partial y^2} \\ & \rho\left(u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y}\right) - \left(u\frac{\partial p}{\partial x} + v\frac{\partial p}{\partial y}\right)N = \\ & \qquad \frac{\mu}{\mathrm{Re}_* \mathrm{Pr}} \cdot \frac{\partial^2 T}{\partial y^2} + \frac{N\mu}{\mathrm{Re}_*}\left[\frac{4}{3}\left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right] \end{aligned}\right\} \tag{1.40}$$

令 $Z' = (\rho, u, v, T, y_u, y_v, y_T)^T$，其中 $y_u = \frac{\partial u}{\partial y}, y_v = \frac{\partial v}{\partial y}, y_T = \frac{\partial T}{\partial y}$，则(1.40)可化为：

$$A''\frac{\partial Z'}{\partial x} + B''\frac{\partial Z'}{\partial y} = H' \tag{1.41}$$

$$A'' = \begin{bmatrix} u & \rho & 0 & 0 & 0 & 0 & 0 \\ \dfrac{T}{\rho} & u & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & u & 0 & 0 & 0 & 0 \\ -TuN & 0 & 0 & cu-\rho u & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B'' = \begin{bmatrix} v & 0 & \rho & 0 & 0 & 0 & 0 \\ 0 & v & 0 & 0 & -d & 0 & 0 \\ \dfrac{T}{\rho} & 0 & v & 1 & 0 & -\dfrac{4d}{3} & 0 \\ -TvN & 0 & 0 & cv-\rho v & 0 & 0 & -\omega \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$H' = \left(0,0,0,\frac{\mu}{\text{Re}_*}\left(\frac{4}{3}y_v^2 + y_u^2\right),y_u,y_v,y_T\right)^T$$

式中$c=\rho N, d=\dfrac{N\mu}{\rho\text{Re}_*}, \omega=\dfrac{\mu}{\text{PrRe}_*}$。

(1.41)的特征方程为$\det(\sigma_1 a_{ij}'' + \sigma_2 b_{ij}'')=0$,解得:

$$(u\sigma_1 + v\sigma_2)\sigma_2^6 = 0$$

$$\sigma_2^6 = 0,\ \lambda_7 = -\frac{\sigma_1}{\sigma_2} = \frac{v}{u} \tag{1.42}$$

可见,除质量方程外,方程组(1.40)是抛物型的。

这里必须对次特征的问题加以说明。前面讲过,除了所讨论的方程组的主特征,即由其最高阶偏导数前系数来确定的方程性质外,必须将最高阶导数(即耗散项)略去后的方程的类型加以分析。对于定常的N.S.方程组(1.31)因已呈椭圆型的特征,就不必再考虑次特征。Euler方程(1.38)本身的特征即表示其性质,没有次特征的问题。但抛物型的N.S.方程(1.40)虽呈抛物型方程的性质,其次特征则是由Euler方程组来确定的。因此当亚声速流场由(1.40)来计算时,次特征出现2个复数根,从而具有椭圆型的性质,即流场中的沿主流方向的压力梯度,可以由扰动在亚声速流场中逆主流方向前传,使全流场互相影响。若用沿主流方向空间步进求解而不在计算方法上作些变化,将出现不适定的问题。超声速流场中(1.40)是抛物一双曲型方程组,不必在计算方法上采取其他措施。

4. 不可压定常N.S.方程组与Euler方程组的数学性质

对二维问题,无因次形式不可压定常N.S.方程组为:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \tag{1.43a}$$

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{\partial p}{\partial x} = \frac{1}{\text{Re}}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \tag{1.43b}$$

$$u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{\partial p}{\partial y}=\frac{1}{\mathrm{Re}}\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right) \tag{1.43c}$$

对二维定常流可引入流函数 $\Psi(x,y)$ 及旋度的概念 $\omega=\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}$。对(1.43b)的式取 $\frac{\partial}{\partial y}$，对(1.43c)取 $\frac{\partial}{\partial x}$，然后两式相减得：

$$\Psi_y\omega_x-\Psi_x\omega_y=\frac{1}{\mathrm{Re}}(\omega_{xx}+\omega_{yy}) \tag{1.44a}$$

由 ω 的定义得：

$$\Psi_{xx}+\Psi_{yy}=-\omega \tag{1.44b}$$

按(1.23)式，(1.44)两方程的 $\Delta<0$，因而(1.43)是典型的椭圆型方程，可以提边值问题。但如果雷诺数很大，将(1.43)化为不可压流的 Euler 方程：

$$\left.\begin{aligned}&\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0\\&u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{\partial p}{\partial x}=0\\&u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{\partial p}{\partial y}=0\end{aligned}\right\} \tag{1.45}$$

同理，(1.44)变为：

$$\left.\begin{aligned}&\Psi_y\omega_x-\Psi_x\omega_y=0\\&\Psi_{xx}+\Psi_{yy}=-\omega\end{aligned}\right\} \tag{1.46}$$

因为(1.46)的第一式主要变量是 ω，正是对 ω 的双曲型方程，因此，(1.45)是椭圆一双曲型混合方程组，方程组的类型是不确定的。

5. 其他定常流动的简化方程的数学性质

(1) 全位势方程

$$(a^2-\phi_x^2)\phi_{xx}-2\phi_x\phi_y\phi_{xy}+(a^2-\phi_y^2)\phi_{yy}=0 \tag{1.10}$$

$$\Delta=(-\phi_x\phi_y)^2-(a^2-\phi_x^2)\cdot(a^2-\phi_y^2)=a^2(q^2-a^2) \tag{1.47}$$

$q=u^2+v^2=\phi_x^2+\phi_y^2$ 表示合速度。由(1.47)可见：超声速流动(1.10)是双曲型的；亚声速流动(1.10)是椭圆型的。

(2) 小扰动速势方程

$$(1-M_\infty^2)\phi'_{xx}+\phi'_{yy}=0 \tag{1.11}$$

$$\Delta=(M_\infty^2-1)$$

因此在亚声速流动中是椭圆型的；超声速流动中是双曲型的。

(3) 跨声速流动的小扰动速势方程

$$\left(1-M_\infty^2-\frac{\gamma+1}{u_\infty}M_\infty^2\phi'_x\right)\phi'_{xx}+\phi'_{yy}=0 \tag{1.12}$$

因为流场中同时存在亚声速流区和超声速流区，情况较为复杂。可以证明，当 M_∞^2 接近于 1 时，$\Delta=-\left(1-M_\infty^2-\frac{\gamma+1}{u_\infty}M_\infty^2\phi'_x\right)\approx-(1-M^2)$，此 M 即当地马赫数。因此当流场中一部分流区 $M<1$ 时，流动控制方程是椭圆型的；$M>1$ 的流区方程是双曲型的；交界线

$M=1$ 处方程呈抛物型。所以，求解全流场就遇到所谓混合型控制方程的问题。此混合型方程的解是存在的，这已由 Tricomi 所证明。下面简单介绍一下 Tricomi 定解问题的提法。

如图 1.1，ADB 包围的是亚声速区，AB 为交界线，方程呈椭圆型；ACB 为超声速区，方程转为双曲型。Tricomi 定解问题的提法：在 ADB 及 AC 上给定适当光滑的函数，求出方程的解。此解必须在整个解区内二阶导数连续且满足定解条件。在交界线 AB 上一阶导数连续且在 A、B 点附近允许解的参变量有一定的奇性。

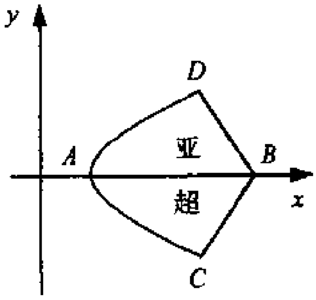


图 1.1　混合型问题示意图

求解跨声速流场时，ADB 边界上边界条件必须给出，AC 上的初始条件也给出。设在交界线 AB 上的变量 $\phi'(x)|_{AB}$ 适当给定，则在 ADBA 区可提 Dirichlet 边值问题，对椭圆型方程是适定的。在 ACBA 区内提初边值问题，对双曲型方程也是适定的，只要在 AB 上适当给定 $\phi(x)|_{AB}$ 就行了。Tricomi 提出了 $\phi'(x)|_{AB}$ 的给定方法。

6. 非定常条件下各方程的数学类型

(1) 非定常流 Euler 方程

一维非定常 Euler 方程可由(1.38)式加上时间项，并去掉 y 向导数项得：

$$\left.\begin{aligned}&\frac{\partial\rho}{\partial t}+\rho\frac{\partial u}{\partial x}+u\frac{\partial\rho}{\partial x}=0\\&\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+\frac{1}{\rho}\frac{\partial p}{\partial x}=0\\&\frac{\partial p}{\partial t}+u\frac{\partial p}{\partial x}+a^2\rho\frac{\partial u}{\partial x}=0\end{aligned}\right\}\tag{1.48}$$

写成列矢量形式：

$$\frac{\partial f}{\partial t}+A\frac{\partial f}{\partial x}=o\tag{1.49}$$

其中，

$$f=(\rho,u,p)^T,\ A=\begin{pmatrix}u&\rho&0\\0&u&\frac{1}{\rho}\\0&a^2\rho&u\end{pmatrix}$$

(1.49)的特征方程为 $|A-I\lambda|=0$，即：

$$(u-\lambda)^3-a^2(u-\lambda)=0$$

$$\lambda_1=u,\ \lambda_{2,3}=u\pm a\tag{1.50}$$

可见，一维非定常 Eluer 方程的三个特征值为相异实数，所以，无论是超声速或亚声速流动，方程均是双曲型的，其特征线斜率分别是 $\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)_{1,2,3}=(u,u-a,u+a)^T$。

二维非定常流的方程形式，由(1.38)加上时间项得：

$$\left.\begin{aligned}&\frac{\partial\rho}{\partial t}+\rho\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+u\frac{\partial\rho}{\partial x}+v\frac{\partial\rho}{\partial y}=0\\&\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial x}=0\\&\frac{\partial v}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial v}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial y}=0\\&\frac{\partial p}{\partial t}+u\frac{\partial p}{\partial x}+v\frac{\partial p}{\partial y}+a^2\rho\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)=0\end{aligned}\right\}\tag{1.51a}$$

写成列矢量形式：

$$\frac{\partial U}{\partial t}+A'\frac{\partial U}{\partial x}+B'\frac{\partial U}{\partial y}=0\tag{1.51b}$$

A',B'分别见(1.39)式。

对二维非定常流，没有明显的特征线可言，判定方程的性质通常是这样进行的：

由控制方程 $A_0\frac{\partial U}{\partial t}+A\frac{\partial U}{\partial x}+B\frac{\partial U}{\partial y}=0$ 中的系数矩阵 A_0,A,B 列出特征方程 $\det[a_{0ij}\sigma_1+a_{ij}\sigma_2+b_{ij}\sigma_3]$，解出的 $\sigma_1,\sigma_2,\sigma_3$ 表示特征面法线的三个方向导数；如果此法线是存在的，那么其三个方向导数必然是实数，方程组是双曲型的。换句话说，算出特征法线存在，垂直于它们的特征面也就存在，方程组的双曲性就确定了。

对方程(1.51)，其特征法线方向导数的特征方程为：

$$\det[\sigma_1\delta_{ij}+\sigma_2a'_{ij}+\sigma_3b'_{ij}]=0$$

即：

$$(\sigma_1+u\sigma_2+v\sigma_3)^2[(\sigma_1+u\sigma_2+v\sigma_3)^2-a^2(\sigma_2^2+\sigma_3^2)]=0\tag{1.52a}$$

$$(\sigma_1+u\sigma_2+v\sigma_3)_{1,2}=0\tag{1.52b}$$

$$(\sigma_1+u\sigma_2+v\sigma_3)_{3,4}=\pm a\sqrt{\sigma_2^2+\sigma_3^2}\tag{1.52c}$$

设法线矢量为 $\boldsymbol{N}=\sigma_1\boldsymbol{i}+\sigma_2\boldsymbol{j}+\sigma_3\boldsymbol{k}$，伪速度矢量为 $\boldsymbol{V}=1\boldsymbol{i}+u\boldsymbol{j}+v\boldsymbol{k}$，则(1.52b)可表示为 $\boldsymbol{N}\cdot\boldsymbol{V}=0$。

由于 $\boldsymbol{N}$ 的长短是任意的，故可令物理平面 xy 面上特征法线的长度为 $l=\sqrt{\sigma_2^2+\sigma_3^2}$。以此为准，(1.52c)可表示为：

$$n_t+un_x+vn_y=\boldsymbol{n}\cdot\boldsymbol{V}=\pm a$$

从面(1.52a)表示为

$$\begin{cases}\boldsymbol{n}\cdot\boldsymbol{V}=0\\\boldsymbol{n}\cdot\boldsymbol{V}=\pm a\end{cases}\tag{1.53}$$

其中

$$n_x=\frac{\sigma_2}{\sqrt{\sigma_2^2+\sigma_3^2}},\ n_y=\frac{\sigma_3}{\sqrt{\sigma_2^2+\sigma_3^2}},\ n_t=\frac{\sigma_1}{\sqrt{\sigma_2^2+\sigma_3^2}}$$

由此可见，特征法线是存在的，其方向导数为实数，因此特征面也是存在的。故二维非定常 Euler 方程组是双曲型的。

应当说明，特征面并非总是存在的。如对定常三维可压流 Euler 方程的分析表明，只

有超声速流动中才有特征面存在。因此三维定常 Euler 方程在亚声速流场中也是椭圆型的。

(2) 非定常可压流 N. S 方程组

在(1.33)前面加上 ρ,u,v,T 的时间导数项得：

$$A_0\frac{\partial Z}{\partial t}+A\frac{\partial Z}{\partial x}+B\frac{\partial Z}{\partial y}=H \tag{1.54}$$

其中 Z,A,B,H 与(1.34)的一样，

$$A_0=\begin{bmatrix}1&0&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0&0\\0&0&0&0&1&0&0&0&0&0\\0&0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0\end{bmatrix}$$

此方程组特征法线的方程是：

$$\det[\sigma_0 a_{0ij}+\sigma_1 a_{ij}+\sigma_2 b_{ij}]=0$$

通过运算可得

$$\begin{cases}\sigma_0+u\sigma_1+v\sigma_2=0\\ \sigma_0^6=0\\ \sigma_0^3+\sigma_1^3+\sigma_2^3=0\end{cases}$$

对于 σ_0 来说，除有一个存在的实根可用 $\boldsymbol{N}\cdot\boldsymbol{V}=0$ 表示外，其他均是实重根。因此可以说非定常可压流 N. S. 方程是抛物－双曲型方程，整体呈抛物型。这个结论十分重要，也比较易于接受。从数理方程已知，$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$或$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}$是典型的抛物型方程，面N. S. 方程的性质主要是由其最高阶空间导数项所确定，因此对非定常可压流 N. S. 方程中除时间导数项外就是粘性耗散项中的二阶导数项决定了方程的类型。其他如对粘性耗散项中略去某些小项的薄层方程也具有抛物型方程的性质。

(3) 非定常不可压 Euler 方程

由(1.45)式，加上 u,v 的时间导数项得：

$$\left.\begin{aligned}&\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0\\&\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{\partial p}{\partial x}=0\\&\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{\partial p}{\partial y}=0\end{aligned}\right\} \tag{1.55a}$$

写成列矢量形式为：

$$A_0'U_t + A'U_x + B'U_y = 0 \tag{1.55b}$$

其中：

$$U = (u,v,p)^T, A_0' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A' = \begin{bmatrix} u & 0 & 1 \\ 0 & u & 0 \\ 1 & 0 & 0 \end{bmatrix}, B' = \begin{bmatrix} v & 0 & 0 \\ 0 & v & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

方程组特征法线的方程是：

$$\det[\sigma_0 a_{0,ij}' + \sigma_1 a_{ij}' + \sigma_2 b_{ij}'] = 0$$

由此解得：

$$(\sigma_0 + u\sigma_1 + v\sigma_2)(\sigma_1^2 + \sigma_2^2) = 0$$

由$(\sigma_0+u\sigma_1+v\sigma_2)=0$得$\boldsymbol{n}\cdot\boldsymbol{V}=0$，存在特征面；由$(\sigma_1^2+\sigma_2^2)=0$知存在虚根。因此方程组(1.55)是双曲—椭圆型的。严格地说，此方程组的数学类型是不确定的。

1.3.3 流体力学控制方程组的数值解法

一、典型偏微分方程的定解条件

为了保证偏微分方程解的适定性，定解条件的提法必须与方程的数学类型相适应。这也正是我们先进行控制方程数学分类的原因。

对椭圆型偏微分方程提边值问题是适定的，且第一、第二、第三类边值问题均可以提，但不允许提初值问题。对抛物和双曲型方程可以提初值问题，也可以提初边值问题，且第一、二、三类边值问题均可以提，但不允许提纯边值问题。双曲型与抛物型方程边值问题的提法有一定的差别，但属于同一类型。不遵守定解条件要与方程类型相适应的规定，解将不适定，数值解可能出现误差的指数增长，直至发散。

二、流体力学控制方程组的分类与求解

现将流体力学控制方程组的性质列表如下，并就定解条件提法，论述求解方法、分析求解时应注意的问题。椭圆型与双曲或抛物型混合的方程组处理起来很困难，必须在数学上进行讨论与分析。根据上节的分析，可将分类列于表1.1。

表1.1 方程组数学分类

	可压 N.S.	可压 Euler	不可压 N.S.	不可压 Euler
定常	椭圆—双曲	亚声速：椭圆—抛物 超声速：双曲—抛物	椭圆	椭圆—双曲
非定常	抛物—双曲	抛物—双曲	椭圆—抛物	椭圆—双曲

从上面解的适定性要求出发，分析定常可压N.S.方程组。亚声速流动的定常Euler方程与不可压Euler方程是椭圆—抛物或椭圆—双曲型方程组，这给定解条件的提法带来了困难。若提初边值问题，椭圆型方程的解不适定；而提纯边值问题，则双曲或抛物型方程的解不适定。这种不定型方程使求解很困难，而定常可压流问题又是工程中经常遇到的问题。

解决上述矛盾的方法是通过一些变换使方程性质变成易求解的。一是所谓时间相关法或稳定法，即将上述定常可压流 N. S. 方程或可压 Euler 方程化为非定常方程，使之由椭圆—双曲混合的不确定方程化为双曲—抛物型方程，提初边值问题而使解适定，并采用时间步进法求解。其实时间相关法(Time-Dependent-Method)模仿了实际的物理现象，如火箭发动机喷管中的定常流场，就是通过起动过程随时间的增长而达到稳定的定常流动的。所以加了时间相关项后，理论上说迭代步长趋于无限大($t\Rightarrow\infty$)，达到定常的解，实际上不必趋于无限大，只要步进后物理量的相对变化小于某一个小量如 10^{-4}，即可认为达到定常状态。这种方法的优点是可充分利用初值问题的各种技巧，得出较好的结果，数学处理与编程较简便。缺点是计算机内存与机时耗费巨大。一般的计算机要解三维实用问题，目前还存在一定困难。另一种办法是将流体力学控制方程中某些次要项略去，或采用其他近似方法，使方程改变类型，达到求解的目的。如 1.2 节所介绍的 N. S. 方程的抛物化，即是略去时间导数项且将沿主流方向的导数在粘性项中全略去，使方程组呈抛物型，然后进行空间步进求解。这是一个简便的方法，但也有不足之处。首先，略去了方程中的某些项，对有些流动，如分离流，有较大的误差；而沿主流方向如有回流，或者是亚声速区次特征的椭圆型，若未加处理即步进求解会使解振荡发散，这些都限制了抛物化方法的使用。当然，若采用全场往复迭代可以解决上述某些问题。这说明在数值方法方面还有大量的工作要做。对于不可压流问题，通常有有限元法、质量方程中加伪压缩项的方法、流函数法等。

三、求解双曲、抛物型方程时解析定解条件的提法

双曲和抛物型方程虽然均可提初边值问题，但边界条件的提法上是有些差异的。下面先从一维非定常可压 Euler 方程(双曲型)出发来分析解析边界条件的提法。

1. 一维非定常流的解析边界条件

设控制方程为：

$$\left.\begin{aligned}&\frac{\partial\rho}{\partial t}+u\frac{\partial\rho}{\partial x}+\rho\frac{\partial u}{\partial x}=0\\&\frac{\partial u}{\partial t}+\frac{a^2}{\rho}\frac{\partial\rho}{\partial x}+u\frac{\partial u}{\partial x}+\frac{1}{\rho}\frac{\partial p}{\partial s}\frac{\partial s}{\partial x}=0\\&\frac{\partial s}{\partial t}+u\frac{\partial s}{\partial x}=0\end{aligned}\right\}\tag{1.56a}$$

写成列矢量形式：

$$\frac{\partial U}{\partial t}+A\frac{\partial U}{\partial x}=0\tag{1.56b}$$

其中：

$$A=\begin{bmatrix}u&\rho&0\\\frac{a^2}{\rho}&u&\frac{1}{\rho}\cdot\frac{\partial p}{\partial s}\\0&0&u\end{bmatrix},U=\begin{bmatrix}\rho\\u\\s\end{bmatrix}$$

(1.56)的特征方程是：

$$\det[\sigma_1\delta_{ij}+\sigma_2a_{ij}]=0$$

由此解得：

$$\lambda = -\frac{\sigma_1}{\sigma_2} = \frac{\mathrm{d}x}{\mathrm{d}t} = \begin{cases} u+a \\ u-a \\ u \end{cases} \quad 即：\begin{cases} \lambda_1 = u+a \\ \lambda_2 = u \\ \lambda_3 = u-a \end{cases} \tag{1.56c}$$

特征线斜率为相异实数，故与(1.48)一样，(1.56)是双曲型方程组。下面分析其初边值条件的提法，尤其是边值问题的提法。将解区示于图1.2的$x-t$坐标上，设L曲线以右为解区。

求解时应给出$t=0$线上的ρ_0, u_0, s_0，即所谓初始条件；同时应给出L上的条件即所谓边界条件。在$x-t$面上不同的位置，L上的u, a值均不一样，而特征线的斜率也不同。现将AB处的特征线斜率的指向与L曲线斜率的关系作于图上。以A点为例，它有两条特征线指向解区。若欲求λ_{2A}右侧某点上A'的各物理量，必须要求边界线L上给出对应的物理量的值，然后沿此特征线用相容性关系式求出。这是解析法求解的过程。但A'点上与λ_{3A}对应方程描述的物理量不必要求边界线L上的值，而只需用上一个时间步处(已经求得)A''对应的物理量沿接近于λ_{3A}方向用相容性关系式解出，即图中A''点的值已求得，不要求边界上给出。如边界上给定边值，则与A''值有矛盾反面使解不合理。因此为了求得A点附近$n+1$时间步处解区的解，只要给出ρ_L, u_L, s_L中的两个，第三个不必给出，也不能给出，其他各点类推。总之，求解双曲型方程组时，边界上特征线指向解区的数目就是边界条件应给出的物理量的数目。这是解物理问题所必要的边界条件称解析边界条件。给多了解不存在，给少了解不唯一。

作为例子，现分析亚声速气流流过图示Laval喷管的一维非定常等熵流场的解析边界条件的提法。见$x-t$图，解区是$t=0, x=0, x=x_B$三条线组成的沿t方向的开域。$t=0$时的初值应全场给出。在某瞬时$t=t_1$时处，三条特征线见图1.3，入口处是亚声速流，只可能有两条特征线指向解区；出口是超声速气流，没有特征线指向解区。因此，除初始条件外，入口处只须给出例如$u_A(t)$，$\rho_A(t)$两个边界条件，而出口处一个边界条件也不允许给出。初始条件的提法比较有规律，凡变量均需给出初始值。

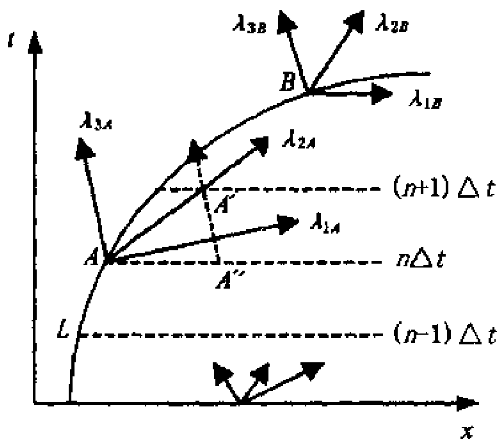


图1.2

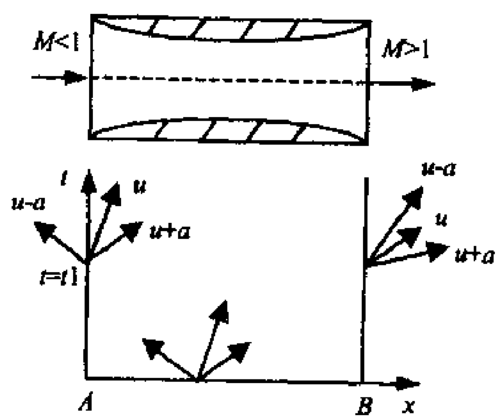


图1.3

2. 二维非定常 Euler 方程解析边界条件的提法

已经分析二维非定常 Euler 方程是双曲型，其控制方程是(1.51)式。当流场的边界及解区是如图 1.4 所示时，解析边界条件的提法为：

(1) 入口 Γ_e^i 处

若垂直于边界处的速度分量 $v_n = \boldsymbol{V}\cdot\boldsymbol{n}$ 是超声速的，则所有变量在入口处要给出，这说明在入口处各特征线均指向解区。若 v_n 是亚声速的，那么必有一条特征线未指向解区，此时通常只给出入口处的熵、总焓与气流流动的方向角，密度可由特征线相容性关系式获得。

(2) 出口 Γ_e^o 处

若出口速度在边界上的法向分量 v_n 为超声速，则不必提边界条件；若 v_n 为亚声速则通常在出口处给定一个边界条件，即压力。

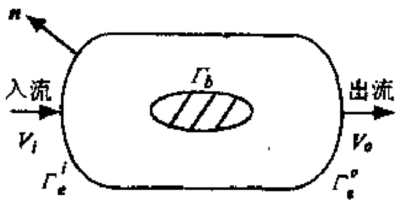


图 1.4

必须指出，对于非定常高雷诺数粘性可压流动，由于求解时入流和出流均远离解区的物面，粘性的影响只在物面附近区域才显得重要，面入流和出流处粘性影响一般不十分严重，因此基本上也可使用上述无粘条件。但若在物面附近取作入、出口条件就不能直接使用上述条件了，所以非定常 N.S. 方程的边界条件要细致分析才行。

3. 抛物型方程的解析定解条件

上面以一维非定常流为基础介绍了双曲型方程解析定解条件的提法，下面讨论(1.56b)式 $\frac{\partial U}{\partial t} + A\frac{\partial U}{\partial x} = 0$。采用当地"冻结"法，$A$ 为常数矩阵。又(1.56)式为双曲型方程组，故 A 必有三个相异的实特征值，存在由右特征矢量组成的相似矩阵 P，使 $\Lambda = P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$，从而(1.56)式可化为

$$P^{-1}\frac{\partial U}{\partial t} + P^{-1}APP^{-1}\frac{\partial U}{\partial x} = 0$$

$$\frac{\partial (P^{-1}U)}{\partial t} + \Lambda\frac{\partial (P^{-1}U)}{\partial x} = 0$$

令 $P^{-1}U = (w1, w2, w3)^T$，则：

$$\left.\begin{aligned} &\frac{\partial w1}{\partial t} + (u+a)\frac{\partial w1}{\partial x} = 0 \\ &\frac{\partial w2}{\partial t} + u\frac{\partial w2}{\partial x} = 0 \\ &\frac{\partial w3}{\partial t} + (u-a)\frac{\partial w3}{\partial x} = 0 \end{aligned}\right\} \tag{1.57}$$

这三个方程均属双曲型方程。最简单的双曲型方程是：

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{1.58}$$

其中 a 是常数，方程特征线的斜率是$\frac{dx}{dt}=a$，以后常用(1.58)的形式来代表最简单的双曲型方程进行讨论。同样可引入：

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2} \tag{1.59}$$

作为最简单的抛物型方程。如何来分析抛物型方程定解条件的提法呢？采用$\frac{\partial^2 u}{\partial x^2}-\varepsilon^2\frac{\partial^2 u}{\partial t^2}=\frac{\partial u}{\partial t}$并取 $\varepsilon^2\Rightarrow 0$ 来接近(1.59)。令 $v=\frac{\partial u}{\partial x}, w=\frac{\partial u}{\partial t}$，则有：

$$\begin{cases}\frac{\partial v}{\partial x}-\varepsilon^2\frac{\partial w}{\partial t}=w\\ \frac{\partial w}{\partial x}-\frac{\partial v}{\partial t}=0\end{cases} \tag{1.60}$$

或写成

$$A\frac{\partial f}{\partial x}+B\frac{\partial f}{\partial t}=S$$

其中

$$f=\begin{pmatrix}v\\w\end{pmatrix}, S=\begin{pmatrix}w\\0\end{pmatrix}, A=\begin{vmatrix}1&0\\0&1\end{vmatrix}, B=\begin{vmatrix}0&-\varepsilon^2\\-1&0\end{vmatrix}$$

特征方程为：$\det(a_{ij}\sigma_1+b_{ij}\sigma_2)=0$，解得：

$$\lambda=\frac{dx}{dt}=-\frac{\sigma_1}{\sigma_2}=\pm\varepsilon \tag{1.61}$$

当 $\varepsilon^2\to 0$ 时，$\varepsilon\to 0$，故(1.59)的特征线是 $dx=0$，见图1.5。此时无论 dx 取任何正负向都可以，因此必须给出初始条件及全部进、出口的边界条件，但在时间前进方向不封顶。这就是抛物型方程解析边界条件的提法。

顺便说明一下，椭圆型方程应当给出所有边界上的边界条件。

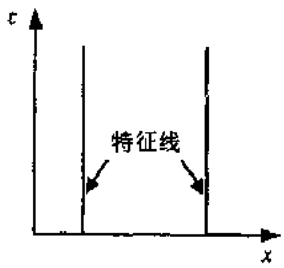


图 1.5

1.4 差分离散的一般原理

1.4.1 选择差分格式的基本要求

为了讨论差分离散与差分格式的构成，我们从最简单的双曲型方程(1.58)及抛物型方程(1.59)出发来进行分析，以弄清其机理。以(1.58)为例，在 $x-t$ 图上，x 方向解区 AE，分成 J 等分，则 $AE=J\cdot\Delta x$；每一个空间间隔点，由 A 到 E，以 $0,1,\cdots,J$ 表示之；某一点 x 处用 $x=j\cdot\Delta x$ 表示。在时间方向用 $t=n\cdot\Delta t$ 表示。通常时间用前差，空间差分可用前差、后差、中心差或其他方法来近似表示。

对(1.58)，时间前差、空间前差的表达方法是：

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}+a\frac{u_{j+1}^n-u_j^n}{\Delta x}=0 \quad (1.62)$$

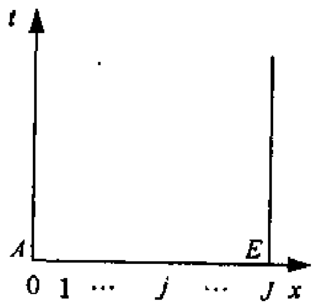


图 1.6

其中：

$$u_{j+1}^n=u[n\Delta t,(j+1)\Delta x]=u(t,x+\Delta x)$$

$$u_j^n=u[n\Delta t,j\Delta x]=u(t,x)$$

所以，(1.62)可写成为：

$$\frac{u(t+\Delta t,x)-u(t,x)}{\Delta t}+a\frac{u(t,x+\Delta x)-u(t,x)}{\Delta x}=0$$

用 FTFS 表示之。其他两种表达式为：

时间前差、空间中心差 FTCS：

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}+a\frac{u_{j+1}^n-u_{j-1}^n}{2\Delta x}=0$$

时间前差、空间后差 FTBS：

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}+a\frac{u_j^n-u_{j-1}^n}{\Delta x}=0$$

其他形式以后使用时再介绍。用差分形式离散方程(1.58)时，似乎 FTFS，FTBS，FTCS 都可以。是不是对一微分方程其差分离散是任意的？回答：不是。对一个确定的微分方程，可以用不同的差分格式离散，但绝不是任意差分离散均可获得合理的解。而且可以使用的差分格式一般也存在着优劣。下面介绍为了取得有意义的微分方程的数值解，差分离散时应当遵循的基本原则。

一、相容性(Consistency)

指的是差分方程必须与微分方程相容，即将差分方程用 Taylor 级数展开后当变量增量 $\Delta t,\Delta x$ 趋于零时，差分方程与微分方程相容。这样才能获得接近于微分方程的解。考虑(1.58)式的 FTBS 差分格式

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}+a\frac{u_j^n-u_{j-1}^n}{\Delta x}=0 \quad (1.63)$$

或

$$u_j^{n+1}=u_j^n(1-\sigma)+\sigma u_{j-1}^n \quad (1.63a)$$

式中 $\sigma=a\dfrac{\Delta t}{\Delta x}$.

对差分方程(1.63)进行 Taylor 级数展开：

$$u_j^{n+1}=u_j^n+\left(\frac{\partial u}{\partial t}\right)_j^n\cdot\Delta t+\frac{1}{2}\cdot\left(\frac{\partial^2 u}{\partial t^2}\right)_j^n\cdot\Delta t^2+\frac{1}{6}\cdot\left(\frac{\partial^3 u}{\partial t^3}\right)_j^n\cdot\Delta t^3+\cdots$$

$$u_{j-1}^n=u_j^n-\left(\frac{\partial u}{\partial x}\right)_j^n\cdot\Delta x+\frac{1}{2}\cdot\left(\frac{\partial^2 u}{\partial x^2}\right)_j^n\cdot\Delta x^2-\frac{1}{6}\cdot\left(\frac{\partial^3 u}{\partial x^3}\right)_j^n\cdot\Delta x^3+\cdots$$

将上面两式代入(1.63)得：

$$\left[\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=\left(-\frac{1}{2}\cdot\frac{\partial^2 u}{\partial t^2}\cdot\Delta t-\frac{1}{6}\cdot\frac{\partial^3 u}{\partial t^3}\cdot\Delta t^2\cdots\right)\right.$$
$$\left.-a\left(\frac{1}{2}\cdot\frac{\partial^2 u}{\partial x^2}\cdot\Delta x-\frac{1}{6}\cdot\frac{\partial^3 u}{\partial x^3}\cdot\Delta x^2+\cdots\right)\right]_j^n \quad (1.64)$$

当 $\Delta x,\Delta t\to 0$ 时，(1.64)式等号右侧趋于零，因此(1.64)与(1.58)相容，于是我们说

用 FTBS 列出的差分方程(1.63)与微分方程(1.58)相容,FTBS 格式解微分方程(1.58)具有相容性。

二、收敛性(Convergence)

设 U 是微分方程的精确解,u 是差分方程的精确解,那么对一个具体问题来说自变量增量趋于零时,u 必须趋近于 U,即:

$$\lim_{\Delta x\to 0,\Delta t\to 0}[U(x,t)-u_j^n]=0 \tag{1.65}$$

式中 $x=j\cdot\Delta x$,$t=n\cdot\Delta t$。满足上式的差分方程的解具有收敛性,否则差分方程的解无意义,不能用它来解对应的微分方程。

三、稳定性(Stability)

用计算机解差分方程时,由于计算中的舍入误差及迭代精度等的限制,计算出来的是近似解,用 $\bar{u}$ 表示。数值计算微分方程的解的误差 ε 可表示为:

$$\varepsilon=U(x,t)-\bar{u}_j^n=(U-u_j^n)+(u_j^n-\bar{u}_j^n) \tag{1.66}$$

称 $\varepsilon_D=U-u_j^n$ 为离散(discretization)误差,称 $\varepsilon_r=u_j^n-\bar{u}_j^n$ 为舍入(round-off)误差。

由此可见:$\lim\limits_{\Delta t\to 0,\Delta x\to 0}\varepsilon_D=0$ 恰好是收敛性定义的表达式;而随着时间迭代步数 n 的增长,ε_r 是否有界则是稳定性问题。如果求解过程不稳定,解可能出现指数增长性振荡,使数值计算无法进行下去。

上面三个要求是用有限差分法(或有限体积法)求解微分方程时选择差分格式必须满足的条件,但并不充分,即并不是满足了上述条件,流体力学控制方程或其他方程即可求得较正确的解了。

四、差分格式的选择必须与微分方程所反映的物理问题的要求相适应

这个方面牵涉的问题较多,且都是十分重要的问题。例如,超声速流场中出现的间断问题,凡满足上述三个要求的差分格式并不是均能较好地捕捉激波,即将激波的位置及跨激波的物理量较好地计算出来。这要求在差分方法上另作一些研究与改进。这方面的问题很多,我们将结合差分方法的研究陆续予以介绍。

1.4.2 Lax 等价定理

在采用一种差分格式写成差分方程并讨论其相容、收敛与稳定的性质时可以看到,相容性是比较容易证明的,只须用 Taylor 级数展开就行;稳定性的证明以后可以介绍诸如 Von Neumann 法等来解决,而收敛性证明一般比较麻烦,但 Lax 等价定理很好地解决了这个难题。

Lax 定理:对给定的适定的线性初值问题,用有限差分方程去近似时,若差分方程满足相容性条件,则其稳定性是其收敛性的充分必要条件。简单地表达为:

适定的线性初值问题,若差分方程相容,则稳定⇔收敛。

这个定律的作用是明显的。对适定的线性初值问题,只须证明其相容性与稳定性,收敛性自然满足,从而省却证明收敛性的麻烦。下面通过引入截断误差等概念来证明 Lax 定律。

一、截断误差(Truncation Error)T_j^n

为了讨论 Lax 等价定律及差分计算中截断误差、离散误差等概念,并便于理解和省

却收敛性证明的复杂性，我们以简单的抛物型方程(1.59)为例。设$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$的精确解是$U(x,t)=U(j\cdot\Delta x,n\cdot\Delta t)$。采用 FTCS 差分格式：

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}=\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{\Delta x^2} \tag{1.67}$$

Taylor 展开得：

$$\left[\frac{\partial u}{\partial t}-\frac{\partial^2 u}{\partial x^2}+\frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2}-\frac{\Delta x^2}{12}\frac{\partial^4 u}{\partial x^4}\right]_j^n+O(\Delta t^2,\Delta x^4)=0$$

式中$O(\Delta t^2,\Delta x^4)$指上式的余项的最大量级是Δt^2与Δx^4。因为$\Delta t,\Delta x$均是较小的数，因此幂次越高，量级越小。将(1.67)减去(1.59)即为截断误差。以$T_j^n(u)$表示：

$$T_j^n(u)=\frac{\Delta t}{2}\left(\frac{\partial^2 u}{\partial t^2}\right)_j^n-\frac{\Delta x^2}{12}\left(\frac{\partial^4 u}{\partial x^4}\right)_j^n+O(\Delta t^2,\Delta x^4) \tag{1.68}$$

较严格的定义：若微分方程为$L_{PD}(U)=0$，L_{PD}是微分方程算子，例如对(1.59)式$L_{PD}=\left(\frac{\partial}{\partial t}-\frac{\partial^2}{\partial x^2}\right)$；差分方程为$L_F(u)=0$，$L_F$是差分方程算子，对(1.67)$L_F=\left(\frac{\partial}{\partial t}-\frac{\partial^2}{\partial x^2}+T\right)_j^n$，则截断误差$T_j^n(u)=L_F(u)-L_{PD}(u)$。由此可见，截断误差是将差分方程减去将差分方程精确解代入微分方程算子而得到的。

应当注意，$L_{PD}(U)=0$，因为U是微分方程的精确解，但$L_{PD}(u)\neq 0$。差分方程的精确解u所以与微分方程精确解不同，是因差分方程中多出$T_j^n(u)$之故。所以差分求解时，实质上是求解加上截断误差项后的微分方程式。显然当自变量$\Delta t,\Delta x$趋于零时，截断误差趋于零，则达到相容性条件。

二、当地截断误差 T_{Lj}^n

为了求得离散误差与某种截断误差的关系，引入当地截断误差。定义为：

$$T_{Lj}^n=L_F(U)-L_{PD}(U)=L_F(U) \tag{1.69}$$

对具体问题(1.59)与差分方程(1.67)，其当地截断误差为：

$$T_{Lj}^n(U)=\left[\frac{\partial U}{\partial t}-\frac{\partial^2 U}{\partial x^2}+T(U)\right]_j^n=T_j^n(U)=\left[\frac{\Delta t}{2}\frac{\partial^2 U}{\partial t^2}-\frac{\Delta x^2}{12}\frac{\partial^4 U}{\partial x^4}\right]_j^n+O(\Delta t^2,\Delta x^4) \tag{1.69a}$$

三、Lax 等价定律的证明

以微分方程(1.59)和差分方程(1.67)为例来证明，只有一定的代表性，但不从一般数学理论证明，比较简便。其证明主要分三步进行：①写出离散误差ε_D与当地截断误差$T_{Lj}^n(u)$的递推关系式；②列出差分方程稳定性的表达式；③证明在满足相容与稳定的条件下，收敛性得到满足，即$\lim\limits_{\Delta t\to 0,\Delta x\to 0}\varepsilon_D=0$。

1. 当地截断误差与离散误差

$$L_F(U)=\frac{U_j^{n+1}-U_j^n}{\Delta t}-\frac{U_{j+1}^n-2U_j^n+U_{j-1}^n}{\Delta x^2}$$

$$L_F(u)=\frac{u_j^{n+1}-u_j^n}{\Delta t}-\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{\Delta x^2}=0$$

$$T_{Lj}^n=L_F(U)=L_F(U)-L_F(u)$$

$$= \frac{1}{\Delta t}\{(U_j^{n+1} - u_j^{n+1}) - (U_j^n - u_j^n) - r[(U_{j+1}^n - u_{j+1}^n)$$

$$- 2(U_j^n - u_j^n) + (U_{j-1}^n - u_{j-1}^n)]\}$$

$$\Delta t \cdot T_{L,j}^n = \varepsilon_{Dj}^{n+1} - [(1-2r)\varepsilon_{Dj}^n + r\varepsilon_{Dj+1}^n + r\varepsilon_{Dj-1}^n] \tag{1.70}$$

其中 $r = \Delta t/\Delta x^2$。将(1.70)写成列矢量形式为：

$$\varepsilon_D^{n+1} = \Delta t \cdot T_L^n + A\varepsilon_D^n + C \tag{1.71}$$

其中

$$\varepsilon_D^n = (\varepsilon_{D1}^n, \varepsilon_{D2}^n, \cdots, \varepsilon_{DJ-1}^n)^T$$

$$C = (r\varepsilon_{D0}^n, 0, \cdots, 0, r\varepsilon_{DJ}^n)^T$$

$$T_L^n = (T_{L1}^n, T_{L2}^n, \cdots, T_{LJ-1}^n)^T$$

$$A = \begin{bmatrix} 1-2r & r & 0 & \cdots & 0 & 0 \\ r & 1-2r & r & \cdots & 0 & 0 \\ 0 & r & 1-2r & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & \cdots & \cdots & r & 1-2r \end{bmatrix}$$

因为边界条件给定，故 $\varepsilon_{D0}^n = \varepsilon_{DJ}^n = 0, C = (0, \cdots, 0)^T$，(1.71)的递推式为：

$$\begin{aligned} \varepsilon_D^{n+1} &= \Delta t \cdot T_L^n + A(\Delta t \cdot T_L^{n-1} + A \cdot \varepsilon_D^{n-1}) \\ &= \Delta t \cdot T_L^n + A \cdot \Delta t \cdot T_L^{n-1} + A^2(\Delta t \cdot T_L^{n-2} + A \cdot \varepsilon_D^{n-2}) \\ &= \cdots \\ &= \Delta t(T_L^n + A \cdot T_L^{n-1} + \cdots + A^n \cdot T_L^0) + A^{n+1} \cdot \varepsilon_D^0 \\ &= \Delta t(T_L^n + A \cdot T_L^{n-1} + \cdots + A^n \cdot T_L^0) \end{aligned} \tag{1.71a}$$

差分格式不同，系数矩阵 A 就不一样。

2. *差分方程的稳定性要求*

差分方程的数值解为：

$$\bar{u}_j^{n+1} = r \cdot \bar{u}_{j-1}^n + (1-2r)\bar{u}_j^n + r \cdot \bar{u}_{j+1}^n$$

这是由于舍入误差存在而得出的近似解。其递推式为：

$$\bar{u}^{n+1} = A \cdot \bar{u}^n = A^{n+1} \cdot \bar{u}^0$$

又差分方程的精确解满足：

$$u^{n+1} = A^{n+1} \cdot u^0$$

稳定性要求舍入误差有界，因为舍入误差是 $J-1$ 维列矢量，所以用无限模，即各列元素中最大值小于一个有限正实数 K 来表示其稳定性：

$$\| \varepsilon_r \|_\infty \leqslant K \tag{1.72}$$

由于 $\| \varepsilon_r^n \|_\infty = \| A^n \varepsilon_r^0 \|_\infty \leqslant \| A^n \|_\infty \| \varepsilon_r^0 \|_\infty = \| A^{n-m} \|_\infty \| \varepsilon_r^m \|_\infty$，通常 $\| \varepsilon_r^0 \|_\infty$ 或某个中间步中 $\| \varepsilon_r^m \|_\infty$ 均为有限值，故稳定性要求(1.72)可表述为：

$$\| A^{n-m} \|_\infty \leqslant K, m = 0, 1, \cdots, n \tag{1.73}$$

3. *相容与稳定条件下必然收敛*

收敛的数学表达式为：

$$\lim_{\Delta t\to 0,\Delta x\to 0}\|\varepsilon_D^{n+1}\|_\infty=0$$

由(1.71a)得：

$$\begin{aligned}\|\varepsilon_D^{n+1}\| &= \Delta t(\|T_L^n\| + \|A\cdot T_L^{n-1}\| + \cdots + \|A^n\cdot T_L^0\|)\\ &\leqslant \Delta t(\|T_L^n\| + K\|T_L^{n-1}\| + \cdots + K\|T_L^0\|)\\ &\leqslant \Delta t\cdot M + n\cdot\Delta t\cdot K\cdot M,\\ M &= \max_{j,n}(T_{Lj}^n)\end{aligned}$$

通常设 $\|T_{Lj}^n\|$ 为有限值，且 $M,\Delta t\cdot n=t,K$ 均为有限值。所以，要使 $\lim\limits_{\Delta t\to 0,\Delta x\to 0}\|\varepsilon_D^{n+1}\|_\infty=0$ 成立，只需 $\lim\limits_{\Delta t\to 0,\Delta x\to 0}M=0$，这正是相容性所要求的。

上面是必要性的证明，充分性的证明留给读者。

1.4.3 显式差分格式与隐式差分格式

对(1.56)差分离散得(1.67)：

$$u_j^{n+1}=u_j^n+r(u_{j+1}^n-2u_j^n+u_{j-1}^n)=ru_{j+1}^n+(1-2r)u_j^n+ru_{j-1}^n$$

即：

$$u^{n+1}=Au^n \tag{1.74}$$

A 就是前面所述的系数矩阵。由(1.74)可见，只要前一时间步的量 u^n 已知，u^{n+1}就能很方便地计算出来。这种格式称为显格式。其一般表达式为：

$$u^{n+1}=Wu^n \tag{1.75}$$

式中 W 是空间差分算子式，由若干个空间差分算子组成。常见的空间差分算子有前差、后差、中心差、平均及移位算子等。它们的含义为：

前差算子$\triangle$：$\triangle u_j=u_{j+1}-u_j=u(x+\Delta x)-u(x)$

后差算子∇：$\nabla u_j=u_j-u_{j-1}=u(x)-u(x-\Delta x)$

中心算子 δ：$\delta u_j=u_{j+\frac{1}{2}}-u_{j-\frac{1}{2}}=u\left(x+\dfrac{\Delta x}{2}\right)-u\left(x-\dfrac{\Delta x}{2}\right)$

(注意：$\delta^2u_j=\delta(\delta u_j)=\delta(u_{j+\frac{1}{2}}-u_{j-\frac{1}{2}})=u_{j+1}-2u_j+u_{j-1}$

平均算子 μ：

$$\mu u_j=\frac{1}{2}(u_{j+\frac{1}{2}}+u_{j-\frac{1}{2}})=\frac{1}{2}\left[u(x+\frac{\Delta x}{2})+u(x-\frac{\Delta x}{2})\right]$$

例如对(1.67)式来说，

$$\begin{aligned}u_j^{n+1}&=u_j^n+r(u_{j+1}^n-2u_j^n+u_{j-1}^n)\\ &=u_j^n+r[(u_{j+1}^n-u_j^n)-(u_j^n-u_{j-1}^n)]\\ &=u_j^n+r[\delta u_{j+\frac{1}{2}}^n-\delta u_{j-\frac{1}{2}}^n]\\ &=u_j^n+r\delta\delta u_j^n\\ &=(1+r\delta^2)u_j^n\end{aligned}\tag{1.76}$$

从而 $W=(1+r\delta^2)$。

式(1.76)或其一般式(1.75)称一步显格式，下一个时间步的变量 u_j^{n+1}(未知量)可直接由各空间点上的已知变量 $u_j^n,u_{j+1}^n,u_{j-1}^n,\cdots$等求出，求解方便简单，但它迭代的步长受到

稳定性的严格限制，因而整个过程需迭代的次数增加了。

差分方程如有形式：

$$Zu^{n+1} = Wu^n \tag{1.77}$$

Z 是另一个作用于未知变量 u_j^{n+1} 上的空间差分算子式。此(1.77)式称为隐式差分格式。例如令 $Z=(1-r\delta^2)$，$W=1$，则有：

$$\frac{u_j^{n+1}-u_j^n}{\Delta t} = \frac{u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1}}{\Delta x^2} \tag{1.78}$$

这是对(1.59)方程的时间前差，并用 $n+1$ 时间步未知量的空间中心差分组成。隐式差分格式计算的时间步长几乎不受稳定性的限制。但每迭代一次所占的时间比显格式要长得多，因为它常常遇到矩阵求逆的问题。以(1.78)为例，求解可写成

$$Au^{n+1} = u^n + C \tag{1.79}$$

其中：

$$A = \begin{bmatrix} 1+2r & -r & 0 & 0 & 0 & 0 & 0 \\ -r & 1+2r & -r & 0 & 0 & 0 & 0 \\ 0 & -r & 1+2r & -r & 0 & 0 & 0 \\ 0 & 0 & \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & -r & 1+2r \end{bmatrix}$$

$u^n=(u_1^n,u_2^n,\cdots,u_{J-1}^n)^T$，$C=(ru_0^{n+1},0,\cdots,0,ru_J^{n+1})^T$，$u_0^{n+1}$，$u_J^{n+1}$ 是边界值，一般是已知的。

一般情况下，要用矩阵求逆来解 u^{n+1}，但对上述问题，A 是三对角矩阵，并且满足对角占优势，故可用简单的追赶法求解。即使如此，也比显式差分格式多费机时。

总之，以(1.77)为准，当 $Z=1$ 时是显式差分格式，$Z\neq$ 常数(即含有算子)时则是隐式差分格式，而 W 的存在对稳定的程度多少有些影响。

1.4.4 差分格式的稳定性

Lax 等价定理指出，适定的线性初值问题的相容性和稳定性是收敛性的充要条件。从前面的证明可见，相容性可用 Taylor 展开而使 Δt，Δx 趋于零，检查截断误差是否趋于零而验证，这较易做到。但收敛性的证明却较为困难。由此可见，用分析的办法提出对稳定性的验证是一个分析差分格式可用性的有效途径。有许多分析稳定性的办法，如能量法、离散扰动法、矩阵法和 Fourier 法等。但应用最广的是 Von Neumann 提出的 Fourier 分析法。

一、不稳定现象

分析微分方程

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{1.58}$$

及其差分格式 FTBS：

$$\frac{u_j^{n+1}-u_j^n}{k} + a\frac{u_j^n-u_{j-1}^n}{h} = 0 \tag{1.63}$$

其中 $k=\Delta t$，$h=\Delta x$，$a>0$。初值 $u_0=f_0(x)$。令 $\sigma=a\dfrac{k}{h}$，则差分方程为：

$$u_j^{n+1} = (1-\sigma)u_j^n + \sigma u_{j-1}^n \tag{1.63a}$$

因为微分方程的解是 $u=f(x-at)$，因此当 $x-at=c$ 时，$f(c)=u=\text{const}$。

取 a,h,σ 为如下值，列出差分方程(1.63a)，进行不同 n 层次的运算，结果见图 1.7。

$a=1,h=0.1,k=0.08,\sigma=0.8$，由(1.63a)知：$u_j^{n+1}=0.2u_j^n+0.8u_{j-1}^n$

$a=1,h=0.1,k=0.1,\sigma=1.0$，由(1.63a)知：$u_j^{n+1}=u_{j-1}^n$

$a=1,h=0.1,k=0.2,\sigma=2.0$，由(1.63a)知：$u_j^{n+1}=-u_j^n+2u_{j-1}^n$

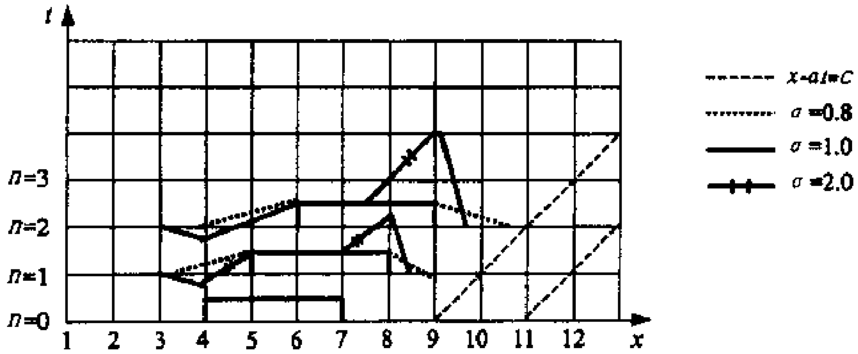


图 1.7

从差分方程的图解(图 1.7)可见，$\sigma=1.0$ 时与微分方程的解一致；$\sigma=0.8$ 时，在解的两端有耗散现象出现；$\sigma=2.0$ 时，解出现振荡且在 $t=n\cdot\Delta t$ 增加时振荡加剧，直到计算失败。这表明，对上述差分格式，$\sigma\leqslant1.0$ 时计算稳定；$\sigma>1.0$ 时计算不稳定，解出现指数增长型的振荡(当迭代时间增加时)。称 σ 为 Courant 数，$\sigma\leqslant1.0$ 为稳定条件。用 FTBS 离散上述方程得到的差分方程是条件稳定的。

二、差分格式的稳定性分析

1. 离散扰动法

从稳定性的定义出发，分析此差分方程的舍入误差随迭代步数的增加是否有界来确定其稳定性。

以方程(1.58)及其差分方程(1.63)为例。

差分方程的精确解：$u_j^{n+1}=(1-\sigma)u_j^n+\sigma u_{j-1}^n$

差分方程的近似解：$\bar{u}_j^{n+1}=(1-\sigma)\bar{u}_j^n+\sigma\bar{u}_{j-1}^n$

从而

$$\varepsilon_{rj}^{n+1}=u_j^{n+1}-\bar{u}_j^{n+1}=(1-\sigma)\varepsilon_{rj}^n+\sigma\varepsilon_{rj-1}^n \tag{1.80}$$

设迭代过程中，在 $t=n\cdot\Delta t, x=j1\cdot\Delta x$ 处开始引入一个舍入误差 $\varepsilon_{rj1}^n=\varepsilon_0^n$，那么在 $t=n\cdot\Delta t$ 处舍入误差的分布情况是：$\begin{cases}\varepsilon_{rj}^n=0, & j\neq j1\\ \varepsilon_{rj}^n=\varepsilon_0^n, & j=j1\end{cases}$，代入(1.80)得

$$\begin{cases}\varepsilon_{rj1}^{n+1}=(1-\sigma)\varepsilon_{rj1}^n=(1-\sigma)\varepsilon_0^n\\ \varepsilon_{rj1+1}^{n+1}=\sigma\varepsilon_{rj1}^n=\sigma\varepsilon_0^n\\ \varepsilon_{rj}^{n+1}=0,\ j\neq j1,\ j1+1\end{cases}$$

为保证求解过程中，差分格式稳定，必须 $|\varepsilon_{rj1}^{n+1}|/|\varepsilon_{rj1}^n|\leqslant1$，即 $|1-\sigma|\leqslant1$，即 $\sigma>0$ 或 σ

$\leqslant 2$。同时，为保证间相摄动而不出现振荡，要求 $\epsilon_{r,j1}^{n+1}/\epsilon_{r,j1}^{n}\geqslant 0$，即 $(1-\sigma)\geqslant 0,\sigma\leqslant 1$。根据 σ 的定义 $\sigma>0$。综合上述要求得(1.63)式稳定的条件：

$$0<\sigma\leqslant 1 \tag{1.82}$$

上述这种分析较直观、概念明确，从中可知，若干差分计算不稳定，舍入误差 ϵ_r 随迭代步数增长不但量值增长很快而且出现振荡，导致求解失败。但它缺乏一般性。

2. Von Neumann 稳定性分析法

(1) Fourier 积分与 Fourier 变换

Fourier 变换是解常系数线性偏微分方程的重要手段之一。将其延伸可用于分析差分方程的稳定性问题。

设 $f(x)$是定义在整个实轴$(-\infty,\infty)$上的实函数，其中任一区间逐段光滑且绝对可积：$\int_{-\infty}^{\infty}|f(x)|^2\mathrm{d}x<\infty$，那么 $f(x)$可用 Fourier 积分来表示：

$$f(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(\beta)\mathrm{e}^{i\beta x}\mathrm{d}\beta$$

其中 $F(\beta)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(\xi)\mathrm{e}^{i\beta\xi}\mathrm{d}\xi$。称 $F(\beta)$为 $f(x)$的 Fourier 变换，$f(x)$为 $F(\beta)$的逆变换。

F 与 f 的关系有如下 Parseval 等式：

$$\|f\|_2=\|F\|_2 \tag{1.83}$$

其中 $\|f\|_2=\left[\int_{-\infty}^{\infty}|f(x)|^2\mathrm{d}x\right]^{\frac{1}{2}}$，$\|F\|_2=\left[\int_{-\infty}^{\infty}|F(\beta)|^2\mathrm{d}\beta\right]^{\frac{1}{2}}$.

(2) 用 Fourier 法分析差分方程的稳定性

以(1.63)为例，

$$u_j^{n+1}=u_j^n-\sigma(u_j^n-u_{j-1}^n),\ \sigma=a\frac{\Delta t}{\Delta x}$$

将上式的定义域延伸到整个实轴上，得：

$$u(x_j,t_{n+1})=u(x_j,t_n)-\sigma[u(x_j,t_n)-u(x_{j-1},t_n)]$$

其中 $x\in(-\infty,\infty)$。将上式视为 $f(x)$，再进行 Fourier 变换得：

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(\beta,t_{n+1})\mathrm{e}^{i\beta x}\mathrm{d}\beta=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(\beta,t_n)\mathrm{e}^{i\beta x}\mathrm{d}\beta$$
$$-\sigma\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(\beta,t_n)\mathrm{e}^{i\beta x}\mathrm{d}\beta-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(\beta,t_n)\mathrm{e}^{i\beta(x-h)}\mathrm{d}\beta\right]$$

即：

$$F(\beta,t_{n+1})=[1-\sigma(1-\mathrm{e}^{-i\beta h})]\cdot F(\beta,t_n)$$

令 $G(h,\beta)=1-\sigma(1-\mathrm{e}^{i\beta h})$，按 Parseval 等式有：

$$\|u(x,t_{n+1})\|_2=\|F(\beta,t_{n+1})\|_2$$
$$\|u(x,t_n)\|_2=\|F(\beta,t_n)\|_2$$

从而

$$\|u(x,t_{n+1})\|_2=\|GF(\beta,t_n)\|_2\leqslant|G|\,\|u(x,t_n)\|_2$$

若 $\bar{u}$ 为差分方程的数值解，考虑了舍入误差的影响，则：

$$\|\bar{u}(x,t_{n+1})\|_2 \leqslant |G|\ \|\bar{u}(x,t_n)\|_2$$

所以

$$\|\varepsilon_r^{n+1}\| \leqslant |G|\ \|\varepsilon_r^n\| = |G^{n+1}|\ \|\varepsilon_r^0\|$$

当 $t=0$，计算开始时存在舍入误差，要使逐次迭代中误差有界，必要求 $\|\varepsilon_r^n\| / \|\varepsilon_r^0\| \leqslant K$，$K$ 为某一确定常数。从而有：

$$|G^n| \leqslant K \tag{1.85}$$

对所有 β 成立。这是差分方程(1.63)稳定的条件，并且对一般初值问题的常系数线性方程均有效，只是 G 的形式不同。

(3) Von Neumann 稳定性准则

从上面的分析可直接导出 Von Neumann 判定初值问题稳定性的条件：$|G^n| \leqslant K$，从而：

$$|G| \leqslant (K)^{\frac{1}{n}} = e^{\frac{1}{n}\ln K} = 1 + \frac{\ln K}{n} + \frac{(\ln K)^2}{2n^2} + \cdots$$

$$= 1 + \frac{k\ln K}{nk} + \frac{k^2(\ln K)^2}{2(nk)^2} + \cdots$$

其中 $nk \leqslant T$，T 是数值计算中总的迭代时间，是有限值，故：

$$|G|_\beta \leqslant 1 + O(k) \tag{1.86}$$

这就是 Von Neumann 判定初值问题差分方程稳定性的准则。可以证明对线性常系数偏微分方程的初值问题，这是充分必要条件。下标 β 表明此不等式对所有 β 均应成立。

为进一步对 Von Neumann 法有所理解，下面对其作些分析。根据 Fourier 变换的说明及不同类型偏微分方程的解由各个分立波迭加而成的说明，可将微分方程的解写成：

$$U(x,t) = \sum_\beta A_\beta e^{-\omega t} e^{i\beta x} = \sum_\beta B_\beta e^{i\beta x} \tag{1.87}$$

式中 A_β 是积分系数；$e^{-\omega t}$ 表示时间变量的影响部分；ω 称圆频率，可以是复数；$\beta=2\pi/\lambda$ 称为波数；λ 是波长。

差分方程的解由于趋近微分方程的解，因此必有：

$$u(x,t) = \sum_\beta A_{1\beta} e^{-\omega nk} e^{i\beta jh} \tag{1.87a}$$

$$\bar{u}(x,t) = \sum_\beta A_{2\beta} e^{-\omega nk} e^{i\beta jh} \tag{1.87b}$$

由于解有可迭加性，对线性方程可取出任一个谐波进行分析而且具有一般性，因此有：

$$\varepsilon_{r\beta}^n = u_j^n - \bar{u}_j^n = A_{r\beta} e^{-\omega nk} \cdot e^{i\beta jh}$$

所以 $|G| = \left|\dfrac{\varepsilon_{r\beta_j}^n}{\varepsilon_{r\beta_j}^{n-1}}\right| = |e^{-\omega k}|$，称为放大因子。

要求 $\lim\limits_{n\to\infty} |\varepsilon_{r\beta j}^n / \varepsilon_{r\beta j}^0| \leqslant K$，即 $|G^n| \leqslant K$，所以 $|G| \leqslant 1 + O(k)$，即(1.86)式。

这告诉我们，在列出差分方程后，可由(1.87)十分方便地将 G 写出并进行“判稳”。

本方法简单方便，工程上常用，但从机理说只限于两层时间变化的初值问题，而且必须是常系数线性方程。事实上，经过冻结系数方法或其他处理，许多非线性的初边值问题

均可用此法进行稳定性分析。而且对一些三层格式，如“$n+1$”、“n”、“$n-1$”，也常借用该法分析。

例 1 微分方程：

$$u_t = u_{xx} + bu \tag{1.88}$$

差分格式：

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2} + bu_j^n$$

其差分方程式为：

$$u_j^{n+1} = r(u_{j+1}^n + u_{j-1}^n) + (1 - 2r)u_j^n + bku_j^n, r = \Delta t/\Delta x^2$$

令 $u_j^{n+1} = Ae^{-\alpha(n+1)k}e^{i\beta jh}$，$u_j^n = Ae^{-\alpha nk}e^{i\beta jh}$。用 $Ae^{-\alpha nk}e^{i\beta jh}$除差分方程式两边得：

$$G = r(e^{i\beta h} + e^{-i\beta h}) + (1 - 2r) + bk$$

$$= 2r\cos\beta h + (1 - 2r) + bk = 1 - 4r\sin^2\frac{\beta h}{2} + bk$$

由 $|G| \leqslant 1 + O(k)$得：

$$\left|1 - 4r\sin^2\frac{\beta h}{2}\right| \leqslant 1, \quad 0 < r \leqslant \frac{1}{2} \tag{1.89}$$

这个例子告诉我们以下三个重要结论：

① 写出差分方程后，放大因子 G 可直接对此方程由 $u_j^n = Ae^{-\alpha nk}e^{i\beta jh}$的方法写出。

② 在进行 G 的分析时可以用代数计算，也可按要求作图。因为 k(即 Δt)是一个小数，通常也可用 $|G| \leqslant 1$ 作为稳定性条件。这个条件应在全部 β 可能的范围内成立。

③ 对于有源项的微分方程，源项是否存在不影响差分方程的稳定性。

例 2 微分方程：

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, a > 0$$

差分方程：

$$u_j^{n+1} = \sigma u_{j-1}^n + (1 - \sigma)u_j^n$$

$$G = (1 - \sigma) + \sigma e^{-i\beta h} = (1 - \sigma) + \sigma\cos\beta h - i\sigma\sin\beta h = G_R + iG_I \tag{1.90}$$

其中 G_R 是实部，G_I 是虚部，可用作图法进行分析。

由 $|G| \leqslant 1$ 得 $G^2 = G_R^2 + G_I^2 \leqslant 1$，而在图 1.8 中，(1.90)式是圆心在$(1-\sigma, 0)$，半径为 σ 的圆，即：$[G_R - (1-\sigma)]^2 + G_I^2 = \sigma^2$。为了保证 $|G| \leqslant 1$，必要求：

$$0 < \sigma \leqslant 1$$

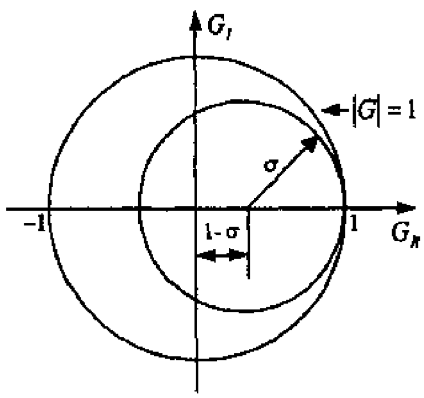


图 1.8

3. 矩阵法分析稳定性

用矩阵法判稳可以考虑边界条件的影响，但数学上处理较为困难。现以简单的抛物型方程为例分析。

微分方程：$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$

定解条件：$u(x,0)=f(x),u(0,t)=u(1,0)=0$

差分格式 FTCS：$u_j^{n+1}=ru_{j-1}^n+(1-2r)u_j^n+ru_{j+1}^n$

写成列矢量形式：$\boldsymbol{u}^{n+1}=A\boldsymbol{u}^n+C$

其中

$$A=\begin{bmatrix} 1-2r & r & 0 & 0 & 0 & 0 & 0 \\ r & 1-2r & r & 0 & 0 & 0 & 0 \\ 0 & r & 1-2r & r & 0 & 0 & 0 \\ 0 & 0 & \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & r & 1-2r \end{bmatrix}_{(J-1)\times(J-1)}$$

$$\boldsymbol{u}^n=(u_1^n,u_2^n,\cdots,u_{J-1}^n)^T$$

$$C=(ru_0^n,0,\cdots,0,ru_J^n)^T=0$$

所以
$$\boldsymbol{u}^{n+1}=A\boldsymbol{u}^n=A^{n+1}\boldsymbol{u}_0$$

同理
$$\bar{\boldsymbol{u}}^{n+1}=A^{n+1}\bar{\boldsymbol{u}}_0$$

从而
$$\boldsymbol{\varepsilon}_r^n=A^n\boldsymbol{\varepsilon}_r^0$$

稳定性要求$\lim\limits_{n\to\infty}\|\boldsymbol{\varepsilon}_r^n\|\leqslant K$。下面通过 A 矩阵的特征值来研究稳定性要求。当 A 是实对称矩阵或其特征值 λ_j 各异时，A 有$(J-1)$个线性独立的特征矢量 $\boldsymbol{v}_j$。这些特征矢量可作为$(J-1)$维向量空间的基，而误差矢量 $\boldsymbol{\varepsilon}_r^0$ 也有$(J-1)$个分量，可以唯一地用 $\boldsymbol{v}_j$ 的线性组合来表示之：

$$\boldsymbol{\varepsilon}_r^0=\sum_{j=1}^{J-1}C_j\boldsymbol{v}_j$$

时间步进 $1\cdot\Delta t$：$\boldsymbol{\varepsilon}_r^1=A\boldsymbol{\varepsilon}_r^0=A\sum\limits_{j=1}^{J-1}C_j\boldsymbol{v}_j=\sum\limits_{j=1}^{J-1}C_jA\boldsymbol{v}_j$

由特征值的定义：$A\boldsymbol{v}_j=\lambda_j\boldsymbol{v}_j$

所以有：

$$\boldsymbol{\varepsilon}_r^1=\sum_{j=1}^{J-1}C_j\lambda_j\boldsymbol{v}_j$$

$$\boldsymbol{\varepsilon}_r^n=\sum_{j=1}^{J-1}C_jA^n\boldsymbol{v}_j=\sum_{j=1}^{J-1}C_j(\lambda_j)^n\boldsymbol{v}_j$$

可见，当 $n\to\infty$，要使$\|\boldsymbol{\varepsilon}_r^n\|\leqslant K$，必要求$|\lambda_j^n|\leqslant K1$。通常只需$\{\max|\lambda_j|\}^n\leqslant K1$，即：

$$\rho(A)\leqslant(K1)^{\frac{1}{n}}=1+\frac{k}{nk}\ln K1+\cdots=1+O(k) \tag{1.91}$$

K，$K1$ 均是有界常数，$\rho(A)$是矩阵 A 的谱半径。此结论与 Von Neumann 法判稳是一致的。通常(1.91)是稳定的必要条件。如果 A 是正规矩阵，则(1.91)也是充分条件。

4. *差分方程组的 Von Neumann 分析法*

微分方程组：

$$\frac{\partial f}{\partial t}+A\frac{\partial f}{\partial x}=0 \tag{1.92}$$

A 是 $m\times m$ 常数矩阵，$f=(f_1,f_2,\cdots,f_m)^T$ 为列矢量。若用差分求解，其显式差分方程组是 $f^{n+1}=Wf^n$。

稳定性的论证过程与 Fourier 法分析一致，可得：

$$\| \varepsilon_r^n \| = \| G^n \| \ \| \varepsilon_r^0 \|$$

稳定性判断准则：

$$\rho(G) \leqslant 1 + O(k) \tag{1.93}$$

即其谱半径 $\rho(G)$ 要满足(1.93)式。式中 G 不是放大因子，而是放大矩阵。

例 3 微分方程：

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

差分方程：

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} = \frac{u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n}{\Delta x^2}$$

或

$$u_j^{n+1} = \frac{2r}{1+2r}(u_{j+1}^n + u_{j-1}^n) + \frac{1-2r}{1+2r}u_j^{n-1}$$

这是时间上的三层格式，可化为两个二层格式求解。令 $v_j^{n+1}=u_j^n$，则

$$u_j^{n+1} = \frac{2r}{1+2r}(u_{j+1}^n + u_{j-1}^n) + \frac{1-2r}{1+2r}v_j^n$$

可写成差分方程组：

$$\begin{pmatrix} u_j^{n+1} \\ v_j^{n+1} \end{pmatrix} = \begin{pmatrix} \dfrac{2r}{1+2r}(e^{i\beta h} + e^{-i\beta h}) & \dfrac{1-2r}{1+2r} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_j^n \\ v_j^n \end{pmatrix}$$

$$G(\beta,k) = \begin{pmatrix} \dfrac{4r\cos\beta h}{1+2r} & \dfrac{1-2r}{1+2r} \\ 1 & 0 \end{pmatrix}$$

稳定性要求 $\rho(G)\leqslant 1+O(k)$，事实上常要求 $\rho(G)\leqslant 1$，只是条件稍苛刻些。下面求 G 的最大特征值。

$$\left(\frac{4r\cos\beta h}{1+2r} - \lambda\right)(0-\lambda) = \frac{1-2r}{1+2r}$$

$$(1+2r)\lambda^2 - (4r\cos\beta h)\lambda - (1-2r) = 0$$

$$\lambda = \frac{2r\cos\beta h \pm \sqrt{1-4r^2\sin^2\beta h}}{1+2r}$$

当 $1-4r^2\sin^2\beta h \geqslant 0$ 时，$|\lambda_{\max}| = \dfrac{2r+1}{1+2r} = 1$；

当 $1-4r^2\sin^2\beta h < 0$ 时，$\lambda = \dfrac{2r\cos\beta h \pm i\sqrt{-1+4r^2\sin^2\beta h}}{1+2r}$，$|\lambda_{\max}| = \left[\dfrac{4r^2-1}{(1+2r)^2}\right]^{\frac{1}{2}} = \left[\dfrac{4r^2-1}{4r^2+4r+1}\right]^{\frac{1}{2}} < 1$，即 $\rho(G)\leqslant 1$。

可见用上述差分格式求解时，对 $r>0$ 的任意值，差分方程是恒稳定的。

第二章　一维抛物型与双曲型方程的差分方法

本章主要叙述一维抛物型与双曲型微分方程差分格式的构成，讨论初值问题的基本处理方法及形成差分格式的机理、求解过程中数值方法上应注意的一些问题等，也涉及到边界条件的处理。因为一维抛物型、双曲型方程比较简单，所以有些数值计算中的方法和机理易于分析清楚。

2.1　一维抛物型方程的显格式

一、一般显格式

一维线性抛物型偏微分方程的一般形式为：

$$\alpha(x,t)\frac{\partial u}{\partial t}=\frac{\partial}{\partial x}\left(a(x,t)\frac{\partial u}{\partial x}\right)+b(x,t)\frac{\partial u}{\partial x}-c(x,t)u \tag{2.1}$$

在某指定的解区内，α，a 为严格的正数，c 是非负数。通常 $t\geqslant 0$，x 可以是无限的或有限的；可提初边值问题，边值问题可提一、二、三类边值。将(2.1)写成：

$$\frac{\partial u}{\partial t}=L(x,t,D,D^2)u \tag{2.2}$$

式中 L 是线性算子，$D=\frac{\partial}{\partial x}$，$D^2=\frac{\partial^2}{\partial x^2}$。对于两层$(n,n+1)$差分格式，我们可以按 Taylor 展开式获得

$$u_j^{n+1}=u(x,t+k)=\left(1+k\frac{\partial}{\partial t}+\frac{k^2}{2!}\cdot\frac{\partial^2}{\partial t^2}+\cdots\right)u_j^n$$

$$=\exp\left(k\frac{\partial}{\partial t}\right)u_j^n=\exp(kL)u_j^n \tag{2.3}$$

同理可得

$$D=\frac{2}{h}\sinh^{-1}\frac{\delta_x}{2}，其中\ \delta_x u_j^n=u_{j+\frac{1}{2}}^n-u_{j-\frac{1}{2}}^n \tag{2.4}$$

从而

$$u_j^{n+1}=\exp\left\{kL\left[jh,nk,\frac{2}{h}\sinh^{-1}\frac{\delta_x}{2},\left(\frac{2}{h}\sinh^{-1}\frac{\delta_x}{2}\right)^2\right]\right\}u_j^n \tag{2.5}$$

解一维抛物型方程所采用的差分格式一般均可归结为(2.5)式。

二、常系数方程的显格式及其精度

微分方程

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$$

此时差分方程由(2.3)式可写成：$u_j^{n+1}=\exp(kL)u_j^n$，$L=D^2$。按(2.4)式展开，

$$D^2 = \frac{1}{h^2}\left(\delta_x^2 - \frac{\delta_x^4}{12} + \frac{\delta_x^6}{90} + \cdots\right)$$

故

$$u_j^{n+1} = \left[1 + \frac{k}{h^2}\delta_x^2 + \frac{k}{2h^2}\left(\frac{k}{h^2} - \frac{1}{6}\right)\delta_x^4 + \cdots\right]u_j^n$$

在目前的问题中，δ_x 均是对 x 取差分。令 $r=k/h^2$，则

$$u_j^{n+1} = \left[1 + r\delta_x^2 + \frac{r}{2}\left(r - \frac{1}{6}\right)\delta_x^4 + \cdots\right]u_j^n \tag{2.6}$$

在(2.6)中，取不同的项，差分格式的精度是不同的。一般定义差分格式当地截断误差的首项确定了格式的精度。如取前两项，则

$$u_j^{n+1} = [1 + r\delta_x^2]u_j^n = u_j^n + r(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \tag{2.7}$$

由(1.69)和(1.69a)知其当地截断误差为：

$$T_{L,j}^n = \left(\frac{k}{2}\cdot\frac{\partial^2 U}{\partial t^2} - \frac{h^2}{12}\cdot\frac{\partial^4 U}{\partial x^4}\right)_j^n + O(k^2,h^4) \tag{2.7a}$$

其中 U 是微分方程的精确解。

当 $\left(\frac{\partial^2 U}{\partial t^2}\right)_j^n$ 和 $\left(\frac{\partial^4 U}{\partial x^4}\right)_j^n$ 为有限值时，(2.7a)的首项是 $O(k,h^2)$ 的量级，故(2.7)式是具有时间一阶精度、空间二阶精度的差分格式。因此，可用截断误差首项中时间与空间增量的量级来确定差分格式的精度。(2.7)式是求解抛物型方程(1.59)常用的差分格式。

如取前三项，则

$$\begin{aligned} u_j^{n+1} &= \left[1 + r\delta_x^2 + \frac{r}{2}\left(r - \frac{1}{6}\right)\delta_x^4\right]u_j^n \\ &= \frac{1}{2}(2 - 5r + 6r^2)u_j^n + \frac{2}{3}r(2 - 3r)(u_{j+1}^n - u_{j-1}^n) - \frac{r}{12}(1 - 6r)(u_{j+2}^n + u_{j-2}^n) \end{aligned}$$

可以证明，其截断误差的首项为 $O(k,h^4)$，因此是时间一阶、空间四阶精度的格式。证明时，认为方程解的各阶导数均存在。这是 5 点格式，计算量增加。但如果设法使 $r=1/6$，则仍保留三点格式而精度提高到与 5 点格式相同。

三、一维自伴型抛物型方程的差分格式

微分方程

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right),\ a(x) \neq 0 \tag{2.8}$$

按(2.2)式可知：$L=aD^2+a'D$。故由(2.3)得

$$u_j^{n+1} = \exp[k(aD^2 + a'D)]u_j^n = [1 + k(aD^2 + a'D) + \cdots]u_j^n$$

令 $D^2u_j^n=\frac{1}{h^2}\delta_x^2u_j^n, Du_j^n=\frac{1}{h}\mu(\delta_xu_j^n)=\frac{1}{2h}(u_{j+1}^n-u_{j-1}^n)$，则

$$u_j^{n+1} = (1 - 2ra)u_j^n + r\left(a + \frac{h}{2}a'\right)u_{j+1}^n + r\left(a - \frac{h}{2}a'\right)u_{j-1}^n \tag{2.9}$$

可以证明此格式之精度是(k,h^2)。

2.2 一维抛物型方程的隐格式

将方程(2.1)的差分格式的一般式(2.3)改写为

$$\exp(-\theta kL)u_j^{n+1}=\exp[(1-\theta)kL]u_j^n \tag{2.10}$$

其中 $0\leqslant\theta\leqslant 1$。这就是不同程度隐式成分的格式。如 $\theta=1$,即为全隐格式;如 $\theta=0$,即为显格式。显格式求解时在未知量所在的时间层"$n+1$"处只有一个网点,而隐格式则有若干个网点处于"$n+1$"层,不能直接求解,对一维抛物型方程常常要解三对角矩阵,对其他问题会更复杂一些。

一、Crank-Nicolson 格式

求解方程(1.59)用(2.10)式,并令 $\theta=1/2$,则有

$$\exp\left(-\frac{kD^2}{2}\right)u_j^{n+1}=\exp\left(\frac{kD^2}{2}\right)u_j^n,\ D^2=\frac{\delta_x^2}{h^2}$$

即

$$\left(1-\frac{k\delta_x^2}{2h^2}\right)u_j^{n+1}=\left(1+\frac{k\delta_x^2}{2h^2}\right)u_j^n \tag{2.11}$$

或

$$(1+r)u_j^{n+1}-\frac{r}{2}(u_{j+1}^{n+1}+u_{j-1}^{n+1})=(1-r)u_j^n+\frac{r}{2}(u_{j+1}^n+u_{j-1}^n)$$

这就是 Crank-Nicolson 格式。它可变为

$$u_j^{n+1}=\frac{r}{2}(u_{j+1}^{n+1}+u_{j-1}^{n+1}-2u_j^{n+1}+u_{j+1}^n-2u_j^n+u_{j-1}^n)+u_j^n$$

即在对二阶导数 $\frac{\partial^2 u}{\partial x^2}$ 差分离散时取了"$n+1$"层和"n"层中心空间差分之平均值。其精度是 (k^2,h^2)。这个格式求解(1.59)不仅在时间精度上比显式提高了,而且其稳定性限制亦大为放宽。事实上可以证明当 $\theta\geqslant 1/2$ 时(2.10)是无条件稳定的格式。

二、差分格式的稳定性分析

令(2.10)中的 $\theta=0$,即得显格式(2.7)式,其相容性由(2.7a)显然满足。下面用 Van Neumann 法判稳:

$$G(k,\beta)=1+r(e^{i\beta h}-2+e^{-i\beta h})=1+2r(\cos\beta h-1)=1-4r\sin^2\frac{\beta h}{2}$$

由 $|G|\leqslant 1$ 得:$-1\leqslant 1-4r\sin^2\frac{\beta h}{2}\leqslant 1$

由右边不等式得:$r>0$;由左边不等式得:$r\leqslant\frac{1}{2}$。所以

$$0<r\leqslant\frac{1}{2} \tag{2.12}$$

可见,用显格式(2.7)计算(2.1)时,h 和 k 不允许任取。当空间步长 h 取定,如果计算中时间步长取得稍大,使 $r>1/2$,$k>h^2/2$,计算会出现振荡导致计算失败。通常 h 是很小的数,如 $h=0.01$,则 $k<0.00005$,因此为了保证稳定性要求,时间步长要取得非常小,达

到一定的时间 T 所需迭代的步数很大。

下面对(2.11)进行稳定性分析：

$$(1+r)G-\frac{r}{2}(Ge^{i\beta h}+Ge^{-i\beta h})=(1-r)+\frac{r}{2}(e^{i\beta h}+e^{-i\beta h})$$

$$G=\frac{(1-r)+r\cos\beta h}{(1+r)-r\cos\beta h}=\frac{1-2r\sin^2\frac{\beta h}{2}}{1+2r\sin^2\frac{\beta h}{2}}$$

显然，只要 $r>0$，都能满足 $|G|<1$，故(2.11)是恒稳的。这是隐格式的好处，但在数值求解时计算每一步要费较多的时间。

三、全隐格式

令(2.10)中的 $\theta=1$，则(1.59)的差分离散式为

$$\exp(-kD^2)u_j^{n+1}=u_j^n,\ D^2=\frac{\delta_x^2}{h^2} \tag{2.13}$$

$$(1-r\delta_x^2)u_j^{n+1}=u_j^n$$

$$u_j^{n+1}-r(u_{j+1}^{n+1}-2u_j^{n+1}+u_{j-1}^{n+1})=u_j^n$$

(2.13)称为全稳式，它在 $r>0$ 时是恒稳的，精度是 (k,h^2)。

四、Du Fort-Frankel 格式

对(1.59)所采用的离散格式中有一种时间三层的 Du Fort-Frankel 格式：

$$\frac{u_j^{n+1}-u_j^{n-1}}{2k}=\frac{1}{h^2}(u_{j+1}^n-u_j^{n+1}-u_j^{n-1}+u_{j-1}^n) \tag{2.14}$$

此格式按稳定性的分析当 $r\geq0$ 时是恒稳的(见 1.4.4 节)。下面分析其相容性。其截断误差为

$$T_j^n=\frac{\Delta t^2}{\Delta x^2}\cdot\left(-\frac{\partial^2u}{\partial t^2}\right)-\frac{\Delta t^2}{6}\cdot\left(-\frac{\partial^3u}{\partial t^3}\right)+\frac{\Delta x^2}{12}\cdot\left(-\frac{\partial^4u}{\partial x^4}\right)+\cdots$$

为保证相容性 $\lim\limits_{\Delta t\to0,\Delta x\to0}T_j^n=0$，要求 $\lim\limits_{\Delta t\to0,\Delta x\to0}\frac{\Delta t^2}{\Delta x^2}=\lim\limits_{\Delta t\to0,\Delta x\to0}\Delta t\cdot\frac{\Delta t}{\Delta x^2}=0$。这说明为达到相容性要求，$\Delta t$ 不能取得太大，要作一定的限制。例如取 Δt 与 Δx 为一个量级，即 $\Delta t/\Delta x=a$，则差分方程与下列微分方程相容：

$$\frac{\partial u}{\partial t}-\frac{\partial^2u}{\partial x^2}+a^2\frac{\partial^2u}{\partial t^2}=0$$

这显然偏离了原方程，因此要求 $\frac{\Delta t}{\Delta x^2}=r$，$r$ 是 1 的量级。相容性要求当 Δx 取一定值时，Δt 的取值更小。此外，格式(2.14)在初始条件和边界条件处理上也有一些困难。

五、非线性抛物型方程

设有非线性方程

$$\frac{\partial u}{\partial t}=\frac{\partial^2(u^\alpha)}{\partial x^2},\alpha\geqslant2 \tag{2.15}$$

为了求解时不受稳定性的限制而采用隐格式：

$$\frac{1}{k}(u_j^{n+1}-u_j^n)=\frac{1}{h^2}[\theta\delta_x^2(u^\alpha)_j^{n+1}+(1-\theta)\delta_x^2(u^\alpha)_j^n]$$

$$(u^a)_j^{n+1} = (u^a)_j^n + k\left(\frac{\partial u^a}{\partial t}\right)_j^n + \frac{k^2}{2}\left(\frac{\partial^2 u^a}{\partial t^2}\right)_j^n + \cdots = (u^a)_j^n + k\left(\frac{\partial u^a}{\partial u}\cdot\frac{\partial u}{\partial t}\right)_j^n + \cdots$$

取

$$(u^a)_j^{n+1} = (u^a)_j^n + k(\alpha u^{a-1})_j^n \cdot \left(\frac{\partial u}{\partial t}\right)_j^n = (u^a)_j^n + \alpha(u^{a-1})_j^n \cdot \Delta u_j,\ \Delta u_j = u_j^{n+1} - u_j^n$$

差分方程变为

$$\frac{1}{k}\Delta u_j = \frac{1}{h^2}\{\theta\delta_x^2[(u^a)_j^n + \alpha(u^{a-1})_j^n\Delta u_j] + (1-\theta)\delta_x^2(u^a)_j^n\}$$

即

$$\begin{aligned}\Delta u_j &= r\{\theta\delta_x^2[(\alpha u^{a-1})_j^n\Delta u_j] + \delta_x^2(u^a)_j^n \\ &= r\theta\alpha[(u^{a-1})_{j+1}^n\Delta u_{j+1} - 2(u^{a-1})_j^n\Delta u_j + (u^{a-1})_{j-1}^n\Delta u_{j-1}] \\ &\quad + r[(u^a)_{j+1}^n - 2(u^a)_j^n + (u^a)_{j-1}^n]\end{aligned} \qquad (2.16)$$

写成纯初值条件下的矩阵式：

$$\begin{pmatrix} 1+2\alpha\theta r(u^{a-1})_1^n & -\alpha\theta r(u^{a-1})_2^n & \cdots & 0 \\ -\alpha\theta r(u^{a-1})_1^n & 1+2\alpha\theta r(u^{a-1})_2^n & -\alpha\theta r(u^{a-1})_3^n & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & -\alpha\theta r(u^{a-1})_{J-2}^n & 1+2\alpha\theta r(u^{a-1})_{J-1}^n \end{pmatrix}\begin{pmatrix}\Delta u_1 \\ \Delta u_2 \\ \vdots \\ \Delta u_{J-1}\end{pmatrix}$$

$$= \begin{pmatrix} -2r(u^a)_1^n & r(u^a)_2^n & \cdots & 0 \\ r(u^a)_1^n & -2r(u^a)_2^n & r(u^a)_3^n & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & r(u^a)_{J-2}^n & -2r(u^a)_{J-1}^n \end{pmatrix}\cdot\begin{pmatrix}u_1^n \\ u_2^n \\ \vdots \\ u_{J-1}^n\end{pmatrix}$$

$$+ \begin{pmatrix} r(u^a)_0^n + \alpha r\theta(u^{a-1})_0^n\Delta u_0 \\ 0 \\ \vdots \\ r(u^a)_J^n + \alpha r\theta(u^{a-1})_J^n\Delta u_J \end{pmatrix}$$

其中 u_0, u_J 是边界值，已给定。Δu 是未知量，可用追赶法求解。

六、三对角矩阵的求解

1. 追赶法

(2.16)的矩阵式可写成如下的递推式：

$$-\eta_j w_{j-1} + \beta_j w_j - \gamma_j w_{j+1} = \delta_j \qquad (2.17)$$

式中 $-\eta_j = -r\theta\alpha(u^{a-1})_{j-1}^n, -\gamma_j = -r\theta\alpha(u^{a-1})_{j+1}^n, -\beta_j = -r\theta\alpha(u^{a-1})_j^n$。$\alpha, \theta, r$ 均是大于零的常数。$w = \Delta u$。

由第一个方程得

$$\beta_1 w_1 - \gamma_1 w_2 = \delta_1,$$

$$\delta_1 = -2r(u^a)_1^n + r(u^a)_2^n + r(u^a)_0^n + \alpha r\theta(u^{a-1})_0^n[u_0^{n+1} - u_0^n]$$

同理

$$-\eta_2 w_1 + \beta_2 w_2 - \gamma_2 w_3 = \delta_2 \qquad (2.18)$$

令 $b_1=\gamma_1/\beta_1, v_1=\delta_1/\beta_1$,则

$$w_1 = b_1 w_2 + v_1$$

将它代入(2.18)得:$w_2=b_2 w_3+v_2$,其中:

$$b_2 = \frac{\gamma_2}{\beta_2 - \eta_2 b_1}, v_2 = \frac{\delta_2 + \eta_2 v_1}{\beta_2 - \eta_2 b_1}$$

类推可得

$$b_0 = 0, \ b_k = \frac{\gamma_k}{\beta_k - \eta_k b_{k-1}}, \ k \geqslant 1,$$

$$v_0 = 0, \ v_k = \frac{\delta_k + \eta_k v_{k-1}}{\beta_k - \eta_k b_{k-1}}, \ k \geqslant 1,$$

$$w_j = b_j w_{j+1} + v_j, \ j = J-1, J-2, \cdots, 1$$

计算时,由于"n"层的数据已知,故可求得 b_k 和 v_k。又由边界条件知 $w_J=u_J^{n+1}-u_J^n$,故可按下列次序求解:

$$w_{J-1} = b_{J-1} w_J + v_{J-1}, \ w_{J-2} = b_{J-2} w_{J-1} + v_{J-2}, \cdots$$

这种方法为追赶法,计算量较小。

2. 对追赶法的分析

上述 $Ax=\delta$ 的线性方程组,A 是三对角矩阵,如果计算稳定,必要求:

① $\eta_j>0, \ \beta_j>0, \ \gamma_j>0$;

② $\beta_j>\eta_j+\gamma_j$;

③ $\beta_j>\eta_{j-1}+\gamma_{j+1}$.

其中①和③保证在回代时每一行对角元素占优势,而①与②保证在消元时每一列对角元素占优势。在(2.16)矩阵式中,上述三个条件均是满足的。这三个条件使追赶法稳定的证明参见 Milchell A. R. 等的著作。

3. 块三对角矩阵的分块追赶法

通常二维或轴对称流动的隐格式常常导致块三对角矩阵式:

$$-A_j u_{j-1} + B_j u_j - C_j u_{j+1} = d_j \tag{2.19}$$

设 $u_J=u_0=0, j=1,2,\cdots,J-1$;$A_j,B_j,C_j$ 为矩阵,u_j,d_j 为列矢量。在二维或轴对称流动问题中,u_j,d_j 是 4 维列矢量,A_j,B_j,C_j 为 4 维矩阵。对(2.19)的求解与前面所述的追赶法完全一样,只是要进行块矩阵的运算。

2.3 抛物型方程的其他问题

2.3.1 抛物型方程的数值边界条件

在分析微分方程类型时,介绍了解析边界条件的提法和为保证方程解的适定性的定解条件的提法。当进行数值解的时候,应考虑数值定解条件的提法。如果提初边值问题,初始条件的提法是将 $t=0$ 时各变量的值给出,但数值边界条件有些特点,现简要归结如下:

① 如果提第二或第三边值问题，其导数项也应用差分表示，此时边界上的差分格式与计算域内点的差分格式可能不一致，因此要考虑边界数值处理对解的精度的影响，且 Von Neumann 判稳法的结果不尽完善（未考虑边界差分离散对稳定性的影响）。

② 差分方程对边界条件的要求与微分方程不一样。例如一维双曲方程，如果有若干特征线不指向解区，那么就对应有若干个边界条件不必提。差分方程的数值边界条件应相应合理给出，不然会导致解的发散。

下面分析边界差分的精度与内点差分格式的精度的关系：

微分方程：$\dfrac{\partial u}{\partial t}=\dfrac{\partial^2 u}{\partial x^2}$

初始条件：$u(x,0)=f(x),x\in[0,\infty)$

边界条件：$\left(\dfrac{\partial u}{\partial x}\right)_{x=0}=C_0(t)$

差分方程取(2.7)式：$u_j^{n+1}=u_j^n+r(u_{j+1}^n-2u_j^n+u_{j-1}^n)$，其精度为$(k,h^2)$。

边界差分格式采用空间前差：

$$\left(\frac{\partial u}{\partial x}\right)_{x=0}=\frac{u_1^{n+1}-u_0^{n+1}}{h}=C_0^{n+1}+O(h) \tag{2.20a}$$

显然，边界上的空间精度低于内点差分解的精度。

为使边界精度与内点精度相匹配，可用截断误差匹配法。

$$\left(\frac{\partial u}{\partial x}\right)_{x=0}=\frac{u_1^n-u_{-1}^n}{2h}=C_0^n+O(h^2)$$

$$u_{-1}^n=u_1^n-2hC_0^n$$

对(2.7)式取 $j=0$，得：

$$\begin{aligned}u_0^{n+1}&=u_0^n+r(u_1^n-2u_0^n+u_{-1}^n)\\&=(1-2r)u_0^n+2r(u_1^n-hC_0^n)\end{aligned} \tag{2.20b}$$

下面举例说明。

考虑 $\dfrac{\partial u}{\partial t}=\dfrac{\partial^2 u}{\partial x^2}$，$0\leqslant x\leqslant 1$

$$u(x,0)=0,$$

$$u(1,t)=0,\left(\frac{\partial u}{\partial t}\right)_{x=0}=-1,\ t>0$$

用分离变量法得解析解：

$$u(x,t)=1-x-\frac{8}{\pi^2}\sum_{\beta=1}^{\infty}(2\beta-1)^{-2}\cos\left[(2\beta-1)\frac{\pi x}{2}\right]\exp\left[-(2\beta-1)\frac{\pi^2 t}{4}\right]$$

由上式可得：

$u(0.4,8)=0.090711$　（$x=0.4,t=8$ 时的解析解）

用差分求解，取 $J=20,h=1/20,r=0.2,k=0.0005$，得：

u(一阶精度处理边界(2.20a)式)$=0.099872,\varepsilon=10\%$

u(截断误差匹配，二阶精度处理边界(2.20b)式)$=0.090468,\varepsilon=0.3\%$

图 2.1 画出了相对误差 ε 随 x 而变化的曲线($t=8.0$)，可见用误差匹配法可使解的

精度提高很多。

数值边界条件的处理较为复杂，对某一方程行之有效的边界条件处理法对另一些方程，尤其是某些非线性方程，可能有问题。常常通过数值试验来解决。

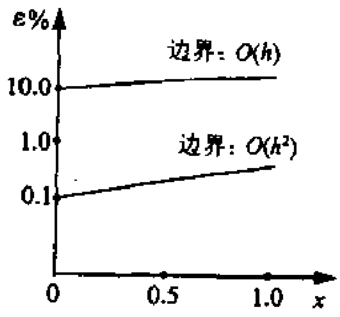


图 2.1

2.3.2 边界条件对稳定性的影响

已知对(1.59)式用差分格式(2.7)，稳定性条件是 $0<r\leqslant 1/2$。在边界上的差分格式是(2.20b)式：$u_0^{n+1}=(1-2r)u_0^n+2r(u_1^n-hC_0^n)$。下面用 Von Neumann 稳定性分析法分析边界上的放大因子 G_0：

$$G_0=(1-2r)+2re^{i\beta h}=1-2r(1-\cos\beta h)+i2r\sin\beta h$$

$$|G_0|=[1-2r(1-\cos\beta h)(8r^2-4r)]^{\frac{1}{2}}$$

由 $|G_0|\leqslant 1$ 得：

$$(8r^2-4r)\leqslant 0,r\leqslant 1/2$$

这个稳定性条件与解区内点的稳定性条件相匹配，因此上述边界条件的处理在精度和稳定性上均与内点匹配。

当边界条件是以导数形式表示，而且此导数与变量有关，那么边界差分处理的稳定性就会对内点的稳定性产生一定的影响。

例如微分方程

$$\begin{aligned}&\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2},0<x<1\\&\frac{\partial u}{\partial x}=A_1(u-B_1),\ x=0,\ t>0\\&\frac{\partial u}{\partial x}=-A_2(u-B_2),\ x=1,\ t>0\end{aligned}\tag{2.21}$$

其中 A_1,A_2,B_1,B_2 均为常数，且 $A_1\geqslant 0,A_2\geqslant 0$。

差分格式 FTCS：$u_j^{n+1}=(1-2r)u_j^n+r(u_{j+1}^n+u_{j-1}^n),j=0,1,\cdots,J$

边界条件采用空间中心差分：

$$u_1^n-u_{-1}^n=2hA_1(u_0^n-B_1),\ x=0$$

$$u_{J+1}^n-u_{J-1}^n=-2hA_2(u_J^n-B_2),\ x=1$$

为便于将边界点和内点一起分析，差分方程写成如下矩阵形式：

$$\begin{pmatrix}u_0\\u_1\\\vdots\\u_{J-1}\\u_J\end{pmatrix}^{n+1}=\begin{pmatrix}1-2r(1+A_1h)&2r&0&0&0\\r&1-2r&r&0&0\\0&\cdots&\cdots&\cdots&\cdots\\\cdots&\cdots&\cdots&\cdots&0\\0&0&\cdots&2r&1-2r(1+A_2h)\end{pmatrix}\begin{pmatrix}u_0\\u_1\\\vdots\\u_{J-1}\\u_J\end{pmatrix}^{n}$$

$$
+\begin{bmatrix} 2rA_1B_1h \\ 0 \\ \vdots \\ 0 \\ 2rA_2B_2h \end{bmatrix}
$$

即：

$$
u^{n+1} = Au^n + C \tag{2.22}
$$

稳定性要求 $\rho(A)\leqslant 1+O(k)$。下面用线性代数中介绍的盖尔圆概念进行分析。

① 方阵 A 的特征值的最大模不能超过任一行或任一列的元素模的和，即特征值在盖尔圆内(Gerschgorin's Theorem)。

② 若 P_s 为第 s 行除对角元素外各元素的模之和，那么矩阵 A 的特征值 λ 至少置于一个圆 $|\lambda-a_{ss}|\leqslant P_s$ 中。a_{ss}是对角元素值。

应分析三类特征值：$j=0$，$j=J$ 及内点，取其中对 r 最严格的限制作为全解区的稳定性限制。

(1) 边界 $j=0$ 处：$a_{ss}=1-2r(1+A_1h)$，$P_s=2r$。

由②得：

$$
|\lambda_0 - [1 - 2r(1 + A_1h)]| \leqslant 2r
$$

加上稳定性条件 $|\lambda_0|\leqslant 1$ 解得：$r_0\leqslant 1/(2+A_1h)$ 与 $r_0\leqslant 1/A_1h$。按最严格的限制应取前者。

(2) 同理，边界 $j=J$ 处：$r_J\leqslant 1/(2+A_2h)$。

(3) 内点处：$r_j\leqslant 1/2$。

所以，全解区的稳定性限制为

$$
r \leqslant \min\left(\frac{1}{2 + A_1h}, \frac{1}{2 + A_2h}\right) \tag{2.23}
$$

2.3.3 解自伴抛物型方程的框元(box)法

这是 Keller 1970 年提出的方法，具有二阶精度、无条件稳定等优点。现介绍如下：

微分方程：

$$
\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right), \quad a(x) \geqslant a_0 \geqslant 0
$$

定解条件：

$$
\begin{cases} u(x,0) = f(x), & 0 < x < 1 \\ u(0,t) = g_0(t), & u(1,t) = g_1(t),\ t > 0 \end{cases}
$$

构造框元法的关键是将二阶方程化为一阶方程组：

$$
\begin{cases} a(x)\dfrac{\partial u}{\partial x} = v \\ \dfrac{\partial v}{\partial x} = \dfrac{\partial u}{\partial t} \end{cases} \tag{2.24}
$$

定解条件：

$$\begin{cases} u(x,0) = f(x),\ v(x,0) = a(x)\dfrac{\mathrm{d}f}{\mathrm{d}x} \\ u(0,t) = g_0(t),\ u(1,t) = g_1(t),\ t > 0 \end{cases}$$

在对方程进行离散时，网格是变间距的，其间距的变化是按要求而选取的，即：

$$x_0 = 0, x_j = x_{j-1} + h_j,\ j = 1,2,\cdots,J$$

$$t_0 = 0,\ t_n = t_{n-1} + k_n,\ n = 1,2,\cdots,N$$

h_j, k_n 是变量，见图 2.2。将方程组按下式离散：

$$a_{j-\frac{1}{2}}\frac{u_j^n - u_{j-1}^n}{h_j} = v_{j-\frac{1}{2}}^n, j = 1,2,\cdots,J$$

$$\frac{1}{h_j}\left(v_j^{n-\frac{1}{2}} - v_{j-1}^{n-\frac{1}{2}}\right) = \frac{1}{k_n}\left(u_{j-\frac{1}{2}}^n - u_{j-\frac{1}{2}}^{n-1}\right),\ n = 1,2,\cdots,N \tag{2.25}$$

式中

$$\begin{cases} \phi_{j\pm\frac{1}{2}}^n = \dfrac{1}{2}(\phi_j^n + \phi_{j\pm1}^n) \\ \phi_j^{n\pm\frac{1}{2}} = \dfrac{1}{2}(\phi_j^n + \phi_j^{n\pm1}) \end{cases}, \phi = u \text{ 或 } v。$$

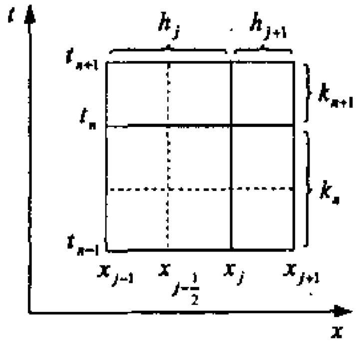


图 2.2

上述格式空间、时间均用后差，系数变量用空间平均值，以保证无条件稳定。它可用于非线性问题。抛物型方程差分格式较单一，大部分用时间前差，空间中心差，用上述知识基本上就可进行分析了。

2.4 一维双曲型方程的差分格式

2.4.1 模型方程

工程问题的控制方程是 Euler 方程及 N.S. 方程。为了反映实际流场，方程是多维的，边界也较复杂。为了弄清数值计算的一些机理性问题，提高计算效率，验证新的格式的效果，可先对简化的模型方程进行研究，然后推广到一般问题中去。

Euler 方程组(非定常)的模型方程是：

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{1.58}$$

或一维双曲型方程组

$$\frac{\partial f}{\partial t} + \frac{\partial F}{\partial x} = \frac{\partial f}{\partial t} + A\frac{\partial f}{\partial x} = 0 \tag{1.49}$$

式中 f 是列矢量，$A = \dfrac{\partial F}{\partial f}$ 是 Jacobian 系数矩阵。

对 N.S. 方程组，最简单的模型方程是 Burgers 方程：

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \bar{\gamma}\frac{\partial^2 u}{\partial x^2} \tag{2.26}$$

这是既有对流项 $u\frac{\partial u}{\partial x}$，又有耗散项 $\bar{\gamma}\frac{\partial^2 u}{\partial x^2}$ 的模型方程，而且在定解条件下有通解，便于与数值解进行比较。式中 $\bar{\gamma}=\mu/\rho$ 是运动粘性系数。

我们先分析(1.58)式的差分解法，再介绍 Burgers 方程的数值解法，尤其对于可压流中间断(激波)的求解作些分析。

2.4.2 双曲型方程的迎风格式

微分方程：

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0 \tag{1.58}$$

如边界条件是周期性的，则可用分离变量法解之，得：

$$u(x,t)=u(j\cdot\triangle x,n\cdot\triangle t)=v(t)e^{i\beta\cdot j\cdot\triangle x}$$

代入(1.58)得：

$$\frac{dv}{dt}e^{i\beta\cdot j\cdot\triangle x}+a\frac{\partial u}{\partial x}=0$$

令

$$\frac{\partial u}{\partial x}=\frac{u_j^n-u_{j-1}^n}{\triangle x},$$

则有：

$$\frac{dv}{dt}e^{i\beta\cdot j\cdot\triangle x}+\frac{a}{\triangle x}(ve^{i\beta\cdot j\cdot\triangle x}-ve^{i\beta(j-1)\cdot\triangle x})=0$$

$$\frac{dv}{dt}=-\frac{a}{\triangle x}(1-e^{-i\beta\cdot\triangle x})v=\alpha v,\ \alpha=-\frac{a}{\triangle x}(1-e^{-i\beta\cdot\triangle x})$$

由此解得：

$$v=Ce^{\alpha\cdot t}$$

为了保证求解过程中时间增加时，$u(x,t)=v(t)e^{i\beta\cdot j\cdot\triangle x}$有界，必须要求 v 有界，即 $\mathrm{Re}(\alpha)<0$(α 的实部小于零)，故必有 $a>0$。所以对于方程(1.58)，如用空间后差离散 $\frac{\partial u}{\partial x}$，则为保证计算稳定，$a>0$；反之，若 $a<0$ 则空间差分必须用前差，否则不稳定。这就是迎风格式，又称逆风格式(Upwinding)。现说明如下：

$\frac{\partial u}{\partial x}$ 前的系数 a 表示波运行的速度，$+a$ 表示正 x 方向的波运行速度。讨论的标准点是 j 点(见图 2.3)。当 $a>0$ 时，因 $\left(\frac{\partial u}{\partial x}\right)_j^n=\frac{u_j^n-u_{j-1}^n}{\triangle x}$，即后差才能使解稳定，差分的指向与波前进的方向相反，故称迎风或逆风。

下面分析迎风格式的稳定性限制与精度。前面一章已讲过，$a>0$ 时(1.58)式的 FTBS 的稳定性条件是 $0<\sigma\leqslant1$，其中 $\sigma=a\frac{\triangle t}{\triangle x}$，故 $\frac{\triangle t}{\triangle x}\leqslant\frac{1}{a}$。而 $\frac{\triangle t}{\triangle x}=\frac{1}{a}$ 正是微分方程(1.58)所对应的特征线的斜率，因此上述稳定性条件可说明为：差分方程的依赖区必须包含微分方程的依赖区。参见图 2.4，分析 $A[(n+1)k,jh]$处的依赖区：对微分方程来说，依赖区是 $BA'A$，差分方程的稳定性要求为 $\frac{\triangle t}{\triangle x}\leqslant\frac{1}{a}$，即依赖区为 $CA'A$。$CA'A$ 包含了 $BA'A$。若

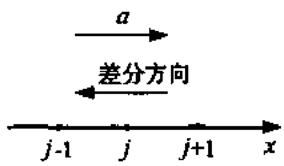


图 2.3

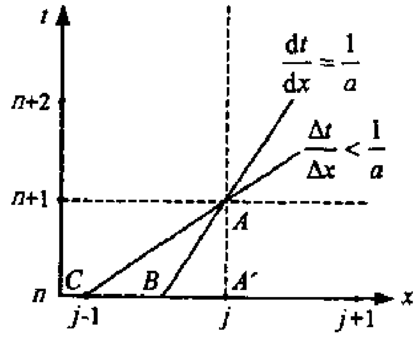


图 2.4

微分方程依赖区大于差分方程的依赖区，初值有某些微小的变化对解的影响不可能在差分方程中反映，而微分方程却受其影响，这样必然使差分方程的解无意义，或不合理。

上述 $\sigma \leqslant 1$ 的条件称 Courant-Friedrichs-Lewy 条件(1928 年提出)，简称 C. F. L. 条件；σ 称 Courant 数(CN)。差分格式的截断误差 $T_j^n = -\left[\frac{\partial^2 u}{\partial t^2} \cdot \frac{\Delta t}{2} + \frac{\partial^2 u}{\partial x^2} \cdot \frac{a\Delta x}{2}\right]_j^n + O(k^2, h^2)$ 。故上式格式为一阶精度的格式(空间与时间均为一阶精度)。

2.4.3 Lax-Wendroff 格式

由于迎风格式是一阶精度，人们为了提高精度曾设想用中心空间差分格式，如对(1.58)用 FTCS 来离散：$\frac{1}{k}(u_j^{n+1} - u_j^n) = -a\frac{1}{2h}(u_{j+1}^n - u_{j-1}^n)$ 。容易证明这是恒不稳定的格式。1954 年 Lax-Friedrichs 提出：

$$\frac{1}{k}\left[u_j^{n+1} - \frac{1}{2}(u_{j+1}^n + u_{j-1}^n)\right] + \frac{a}{2h}(u_{j+1}^n - u_{j-1}^n) = 0 \tag{2.27}$$

由此得

$$G(k,\beta) = \cos\beta h - i\sigma\sin\beta h,\ \sigma = a \cdot \frac{k}{h}$$

$$|G|^2 = |\cos^2\beta h + \sigma^2\sin^2\beta h|^{\frac{1}{2}}$$

由 $|G| \leqslant 1$ 得：$|\sigma| \leqslant 1$，其稳定性限制与迎风格式相同，但其精度未见提高。截断误差的前两项为

$$T_j^n = -\frac{\partial^2 u}{\partial t^2} \cdot \frac{k}{2} + \frac{\partial^2 u}{\partial x^2} \cdot \frac{h^2}{2k} - \frac{\partial^3 u}{\partial x^3} \cdot \frac{ah^2}{6} \tag{2.28}$$

通常 h/k 为一常数，因此 Lax-Friedrichs 格式的精度仍为(h,k)，精度未见提高，而且差分格式选择不如迎风格式那样与特征线走向相联系，使用效果不理想。

为了建立二阶精度的格式，先对 u_j^{n+1} 用 Taylor 级数对 t 展开，使之保持二阶精度，即略去 $O(k^3)$项，然后让空间差分也保持略去 $O(h^3)$项的格式，构成截断误差首项具有 $O(h^2,k^2)$的差分格式。

$$u_j^{n+1} = u_j^n + k\left(\frac{\partial u}{\partial t}\right)_j^n + \frac{k^2}{2}\left(\frac{\partial^2 u}{\partial t^2}\right)_j^n + O(k^3)$$

将

$$\frac{\partial u}{\partial t}=-a\frac{\partial u}{\partial x},\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}$$

代入上式得：

$$u_j^{n+1}=u_j^n-ak\left(\frac{\partial u}{\partial x}\right)_j^n+\frac{a^2k^2}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_j^n+O(k^3)$$

将

$$\left(\frac{\partial u}{\partial x}\right)_j^n=\frac{1}{2h}(u_{j+1}^n-u_{j-1}^n)$$

及

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_j^n=\frac{1}{h^2}(u_{j+1}^n-2u_j^n+u_{j-1}^n)$$

代入上式得：

$$u_j^{n+1}=u_j^n-\frac{\sigma}{2}(u_{j+1}^n-u_{j-1}^n)+\frac{\sigma^2}{2}(u_{j+1}^n-2u_j^n+u_{j-1}^n) \tag{2.29}$$

称此格式为 Lax-Wendroff 格式，并可求得 $G(k,\beta)$ 为

$$G=1-\frac{\sigma}{2}(e^{i\beta h}-e^{-i\beta h})+\frac{\sigma^2}{2}(e^{i\beta h}-2+e^{-i\beta h})$$

要使 $|G|\leqslant 1$，必须 $\sigma\leqslant 1$。所以，Lax-Wendroff 差分格式的稳定性条件与迎风格式一样，但精度提高，其截断误差为

$$T_j^n=\left[\frac{\partial^3 u}{\partial t^3}\cdot\frac{k^2}{6}+\frac{\partial^3 u}{\partial x^3}\cdot\frac{ah^2}{3}\right]_j^n+O(k^3,h^3) \tag{2.30}$$

可见，它具有二阶时间和空间精度。

2.4.4 由特征线构造二阶迎风格式

双曲型方程中有特征线存在。若差分格式的构成符合特征线原理就易于使用。在(1.58)中设 $a>0$，则特征线斜率 $dt/dx=1/a>0$。

如图 2.5 所示，t_n 时各变量 $u_{j-2},u_{j-1},u_j,\cdots$ 均已知，现欲求 u_j^{n+1}，即 P 点的值。过 P 作特征线 PQ，并假设 C.F.L. 条件成立，则由相容性条件有 $u(P)=u(Q)$，且

$$AQ=ak=k\text{ctg}(\theta)。$$

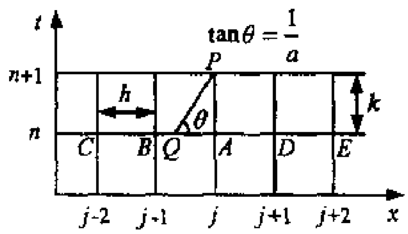


图 2.5

(1) 如在 A、B 两点间线性插值，以确定 Q 处的 u 值，则

$$u(P)=u(Q)=u(B)\frac{AQ}{h}+u(A)\frac{h-AQ}{h}$$

即

$$u_j^{n+1}=\frac{ak}{h}u_{j-1}^n+\left(1-\frac{ak}{h}\right)u_j^n=(1-\sigma)u_j^n+\sigma u_{j-1}^n$$

这就是迎风差分格式(1.63)。

(2) 用 A、B、D 三点抛物线插值得

$$u(P)=u(Q)=u(A)-\sigma[u(A)-u(B)]-\frac{\sigma(1-\sigma)}{2}[u(B)-2u(A)+u(D)]$$

即

$$u_j^{n+1}=u_j^n-\frac{\sigma}{2}(u_{j+1}^n-u_{j-1}^n)+\frac{\sigma^2}{2}(u_{j+1}^n-2u_j^n+u_{j-1}^n)$$

这就是 Lax-Wendroff 格式(2.29)。

(3) 用 A、B、C 三点抛物线插值得

$$u(P)=u(Q)=u(A)-\sigma[u(A)-u(B)]-\frac{\sigma(1-\sigma)}{2}[u(B)-2u(A)+u(C)]$$

即

$$u_j^{n+1}=u_j^n-\sigma(u_j^n-u_{j-1}^n)-\frac{\sigma(1-\sigma)}{2}(u_j^n-2u_{j-1}^n+u_{j-2}^n) \tag{2.31}$$

这就是显式二阶迎风格式。其放大因子为

$$G=1-\sigma(1-\mathrm{e}^{-i\beta h})+\frac{\sigma}{2}(1-\sigma)(1-2\mathrm{e}^{i\beta h}+\mathrm{e}^{-2i\beta h})$$

$$=1-2\sigma\sin^2\frac{\beta h}{2}-\frac{\sigma}{2}(1-\sigma)(4\sin^4\frac{\beta h}{2}-\sin^2\beta h)-i\sigma\sin\beta h\left[1+(1-\sigma)\sin^2\frac{\beta h}{2}\right]$$

$$|G|^2=1-4\sigma(1-\sigma)^2(2-\sigma)\left(\sin^4\frac{\beta h}{2}\right)$$

要使 $|G|\leqslant 1$,必须 $\sigma\leqslant 2$。

可见,此格式稳定性限制放宽,但计算量有所增加。可以证明,其截断误差首项的量级为。$O(k,h^2)$。

2.4.5 隐式格式

1. 隐式 Wendroff 格式

对方程(1.58),如用显式 FTCS 格式离散,其差分方程是恒不稳定的。现用隐式 Wendroff格式

$$\frac{1}{2k}[(u_{j+1}^{n+1}+u_j^{n+1})-(u_{j+1}^n+u_j^n)]+\frac{a}{2h}[(u_{j+1}^{n+1}+u_{j+1}^n)-(u_j^{n+1}+u_j^n)]=0 \tag{2.32}$$

2. Crank—Nicolson 型格式

$$\frac{1}{k}(u_j^{n+1}-u_j^n)+\frac{a}{4h}(u_{j+1}^{n+1}-u_{j-1}^{n+1}+u_{j+1}^n-u_{j-1}^n)=0 \tag{2.33}$$

这些是无条件稳定的格式。格式 2 的截断误差为

$$T_j^n=\left[\frac{k^2}{24}\cdot\frac{\partial^3 u}{\partial t^3}+\frac{ah^2}{6}\cdot\frac{\partial^3 u}{\partial x^3}+\cdots\right]_j^n+O(k^3,h^3)$$

故它为时间与空间均为二阶精度的格式。

由于双曲型方程大都用在解可压流问题,且大都有特征线存在,因此在上面介绍的格式中迎风(一阶与二阶)及 Lax-Wendroff 格式的改型 MacCormack 是在工程上应用较多

的。一些新型格式将在以后介绍。

2.5 一阶双曲型方程组的差分格式

考虑

$$\frac{\partial u}{\partial t}+\frac{\partial F}{\partial x}=0$$

或

$$\frac{\partial u}{\partial t}+A\frac{\partial u}{\partial x}=0 \tag{2.34}$$

式中 $F=F(u)$ 是 u 的组合函数，$u=(u_1,u_2,\cdots,u_m)^T$ 是列矢量。$A=\frac{\partial F}{\partial u}$ 是 Jacobian 矩阵。如果上述方程是双曲型方程组，则 A 为 $m\times m$ 矩阵，且有 m 个线性独立的实特征值，故必存在一个非奇异矩阵 P，使得

$$\Lambda=P^{-1}AP=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & 0 & \cdots\\ \cdots & 0 & \ddots & 0\\ 0 & \cdots & 0 & \lambda_m\end{bmatrix}$$

P 的构成可由 A 求出特征值，再求出相应的右特征矢量而构成矩阵。使用(1.57)式的形式，方程组化为

$$P^{-1}\frac{\partial u}{\partial t}+P^{-1}APP^{-1}\frac{\partial u}{\partial x}=\frac{\partial \omega}{\partial t}+\Lambda\frac{\partial \omega}{\partial x}=0$$

从而化成 m 个非耦合的方程组(式中 $\omega=P^{-1}u$)，其特征值 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 的正负号是不同的。下面针对上述方程构造合适的格式。

一、Lax-Wendroff 格式

将一维双曲型方程的 Lax-wendroff 格式延伸到方程组，得

$$u_j^{n+1}=u_j^n-\frac{\alpha}{2}A(u_{j+1}^n-u_{j-1}^n)+\frac{\alpha^2}{2}A^2(u_{j+1}^n-2u_j^n+u_{j-1}^n)$$

其中 $\alpha=\triangle t/\triangle x$，$u$ 是列矢量，A 是 Jacobian 矩阵，并假设为常数。下面分析其稳定性。

$$G(\beta,h)=I-i\alpha\sin\beta h\cdot A-\alpha^2(1-\cos\beta h)A^2$$

设 A 的特征值为 λ，G 的特征值为 μ，则对矩阵多项式而言，

$$\mu=1-i\alpha\sin\beta h\cdot\lambda-\alpha^2(1-\cos\beta h)\lambda^2$$

$$|\mu|^2=1-4\alpha^2\lambda^2(1-\alpha^2\lambda^2)\sin^4(\beta h/2) \tag{2.35}$$

稳定性要求谱半径 $\rho(G)\leqslant 1$，即 $|\mu|_{\max}\leqslant 1$。由(2.35)得：$\alpha\lambda_{\max}=\alpha\rho(A)\leqslant 1$，从而得 Lax-Wendroff 格式的稳定性条件为

$$\alpha\rho(A)=\frac{\triangle t}{\triangle x}\rho(A)\leqslant 1 \tag{2.36}$$

如果 A 是对称矩阵，则上述条件也是充分条件；否则只是必要条件。证明不再详述。此条件亦称为双曲型方程组的 C.F.L. 条件。

二、一阶迎风格式

考虑 $\frac{\partial u}{\partial t}+A\frac{\partial u}{\partial x}=0$，并将其转化为类似于(1.57)的形式 $\frac{\partial \omega}{\partial t}+\Lambda\frac{\partial \omega}{\partial x}=0$。若有 m 个特征值存在，则有

$$\frac{\partial \omega_i}{\partial t}+\lambda_i\frac{\partial \omega_i}{\partial x}=0,\ i=1,2,\cdots m。$$

通常将迎风格式写成如下形式：

$$\omega_j^{n+1}=\omega_j^n-\alpha\left[\frac{1}{2}(\lambda+|\lambda|)(\omega_j^n-\omega_{j-1}^n)+\frac{1}{2}(\lambda-|\lambda|)(\omega_{j+1}^n-\omega_j^n)\right] \quad (2.37)$$

式中 $\alpha=\triangle t/\triangle x$。此式对于 $\lambda>0,\lambda<0$ 均是迎风的。改写为：

$$\omega_j^{j+1}=\omega_j^n-\frac{\alpha}{2}\lambda(\omega_{j+1}^n-\omega_{j-1}^n)+\frac{\alpha}{2}|\lambda|(\omega_{j+1}^n-2\omega_j^n+\omega_{j-1}^n) \quad (2.38)$$

求出 ω_j^{n+1} 后，由 $u_j^{n+1}=P\omega_j^{n+1}$ 解出 u_j^{n+1}。

下面分析其稳定性：

$$G(\beta,h)=I-i\alpha\sin\beta h\cdot\Lambda+\alpha(\cos\beta h-1)|\Lambda|$$

其特征值为

$$\mu(G)=1-i\alpha\sin\beta h\cdot\lambda_m+\alpha(\cos\beta h-1)|\lambda_m|$$

$$|\mu|^2=[1-2\alpha|\lambda_m|\sin^2(\beta h/2)]^2+\alpha^2\lambda_m^2\sin^2\beta h$$

$$=1-4\alpha|\lambda_m|(1-\alpha|\lambda_m|)\sin^2(\beta h/2)$$

由稳定性要求 $|\mu|_{\max}\leqslant 1$，得 $\alpha\max\limits_m|\lambda_m|\leqslant 1$。因此，$\alpha\max\limits_m|\lambda_m|\leqslant 1$ 满足 Von Neumann 条件。又 G 是对角矩阵，必然是正规矩阵，故上述条件也是 C. F. L. 条件。

三、分离系数矩阵法

前面提到对双曲型方程组若用迎风格式必须将其化成单个非耦合的方程由其特征值的正负来确定格式的空间差分方向，这是较为麻烦的。下面介绍分离系数矩阵法解决此问题。

微分方程组：

$$\frac{\partial u}{\partial t}+A\frac{\partial u}{\partial x}=0$$

u 为 m 维列矢量，A 为 $m\times m$ 矩阵。分离系数矩阵法的求解步骤是：

(1) 对矩阵 A 求特征值 $\lambda_1,\lambda_2,\cdots,\lambda_m$，组成对角矩阵 Λ。

(2) 将正特征值与负特征值分开，形成两个对角矩阵：

$$\Lambda=\Lambda^++\Lambda^-$$

其中 Λ^+ 中的特征值均大于 0，Λ^- 中的特征值均小于 0。凡 Λ^+ 中原 Λ 内特征值为负处均用 0 代替，Λ^- 也同样处理。

(3) 由 A 中各元素分别求出右特征矢量组成矩阵 P，并求 P^{-1}。

(4) 矩阵 $A=P\Lambda P^{-1}=P(\Lambda^++\Lambda^-)P^{-1}=A^++A^-$。式中 A^+ 是特征值为正的系数矩阵；A^- 为特征值为负的系数矩阵。

(5) 将微分方程改写为

$$\frac{\partial u}{\partial t}+A\frac{\partial u}{\partial x}=\frac{\partial u}{\partial t}+A^+\frac{\partial u}{\partial x}+A^-\frac{\partial u}{\partial x}=0 \quad (2.39)$$

(6) 对 $A^+\frac{\partial u}{\partial x}$ 取空间后差，对 $A^-\frac{\partial u}{\partial x}$ 取空间前差，以满足迎风要求。即

$$\frac{\partial u}{\partial t}=-A\frac{\partial u}{\partial x}=-(A^++A^-)\frac{\partial u}{\partial x}=-(A^+u_x^++A^-u_x^-)$$

其中 u_x^+ 要用后差迎风，u_x^- 要用前差迎风。具体求解时为提高精度，常用二步格式，且其中一步要用二阶迎风。具体做法如下：

$$u_j^{\overline{n+1}}=u_j^n-\frac{\Delta t}{\Delta x}[A^+(3u_j^n-4u_{j-1}^n+u_{j-2}^n)+A^-(u_{j+1}^n-u_j^n)]$$

$$u_j^{n+1}=\frac{1}{2}\left\{u_j^n+u_j^{\overline{n+1}}+\frac{\Delta t}{\Delta x}[A^{+\overline{n+1}}(u_j^{\overline{n+1}}-u_{j-1}^{\overline{n+1}})+A^{-\overline{n+1}}(-3u_j^{\overline{n+1}}+4u_{j+1}^{\overline{n+1}}-u_{j+2}^{\overline{n+1}})]\right\}$$

(2.40)

这里说明一下，对 $\frac{\partial u}{\partial x}$ 采用二阶迎风处理的问题，以 $A^+\frac{\partial u}{\partial x}$ 为例，令 $A^+\frac{\partial u}{\partial x}=\frac{A^+(3u_j^n-4u_{j-1}^n+u_{j-2}^n)}{2\Delta x}$，用 Taylor 级数展开可得 $A^+\left(\frac{\partial u}{\partial x}-\frac{\Delta x^2}{3}\frac{\partial^3 u}{\partial x^3}+\cdots\right)_j^n$，因此具有二阶精度。

四、显式多步格式

显式差分格式的一般形式是：$u_j^{n+1}=\tilde{\omega}u_j^n$。$\tilde{\omega}$ 是空间差分算子式。为了提高差分格式的精度，放宽稳定性限制，简化非线性微分方程计算，无论显式或隐式格式常常采用多步格式。一般形式为

$$\begin{cases}u_j^{\overline{n+1}}=\tilde{\omega}^{(1)}u_j^n\\ u_j^{\overline{\overline{n+1}}}=\tilde{\omega}^{(2)}u_j^n+\tilde{\omega}^{(3)}u_j^{\overline{n+1}}\\ \cdots\cdots\end{cases}\tag{2.41}$$

常用的是两步格式。如对双曲型方程组(2.34)的 MacCormack 两步格式：

预测步：

$$u_j^{\overline{n+1}}=u_j^n-\alpha(F_{j+1}^n-F_j^n)$$

校正步：

$$u_j^{\overline{\overline{n+1}}}=u_j^{\overline{n+1}}-\alpha(F_j^{\overline{n+1}}-F_{j-1}^{\overline{n+1}})$$

$$u_j^{n+1}=\frac{u_j^n+u_j^{\overline{\overline{n+1}}}}{2}\tag{2.42}$$

事实上，预测步用空间后差，校正步用空间前差也可写成 MacCormack 两步格式。如果 $\frac{\partial F}{\partial u}=A$，$A$ 设为常数矩阵，则格式(2.42)可写成：

预测步：

$$u_j^{\overline{n+1}}=u_j^n-\alpha A(u_{j+1}^n-u_j^n)$$

校正步：

$$u_j^{\overline{\overline{n+1}}}=u_j^{\overline{n+1}}-\alpha A(u_j^{\overline{n+1}}-u_{j-1}^{\overline{n+1}})$$

$$u_j^{n+1}=\frac{u_j^n+u_j^{\overline{\overline{n+1}}}}{2}\tag{2.43}$$

可以证明，在 A 为常系数矩阵的情况下，上述格式即为 Lax-Wendroff 格式。MacCor-

mack 格式具有时间与空间二阶精度。MacCormack 格式在 20 世纪 70 年代曾得到广泛应用。

当 A 为常数时，方程组(2.34)的 Lax-Wendroff 格式为：

$$u_j^{n+1} = u_j^n - \frac{\alpha A}{2}(u_{j+1}^n - u_{j-1}^n) + \frac{\alpha^2 A^2}{2}(u_{j+1}^n - u_j^n + u_{j-1}^n)$$

上式的两步格式即(2.43)式。

2.6　双曲型方程的修正型方程及解的分析

由于截断误差的存在，差分方程的精确解也不同于微分方程的精确解。下面通过分析截断误差中一些项(包括偶阶和奇阶导数项)对解产生什么影响，进一步了解差分数值求解的机理。

仍以 $\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, a > 0$ 为例，并考虑用 FTBS 离散：

$$u_j^{n+1} = u_j^n - a\alpha(u_j^n - u_{j-1}^n),\ \alpha = \triangle t/\triangle x。$$

将差分方程 Taylor 展开得：

$$\left\{\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = -\frac{\triangle t}{2}\frac{\partial^2 u}{\partial t^2} - \frac{\triangle t^2}{6}\frac{\partial^3 u}{\partial t^3} - \cdots + \frac{a\triangle x}{2}\frac{\partial^2 u}{\partial x^2} - \frac{a\triangle x^2}{6}\frac{\partial^3 u}{\partial x^3} + \cdots\right\}_j^n \tag{2.44}$$

因截断误差项中有时间导数和空间导数项，不便于看清问题的机理。现将各时间导数项均化为空间导数项，导出修正方程。

2.6.1　差分方程的修正型方程(改型方程)

由于在差分方程中差分解不满足 $\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0$ 而满足(2.44)式，因此在改型时可采用如下方法：对差分方程取时间导数并乘上适当的系数消去时间导数项，主要取得改型修正方程截断误差的前两项。

由表 2.1 得

$$\left[\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} + \frac{ka^2 - ah}{2}\cdot\frac{\partial^2 u}{\partial x^2} + \left(\frac{ah^2}{6} + \frac{a^3k^2}{3} - \frac{a^2kh}{2}\right)\cdot\frac{\partial^3 u}{\partial x^3} + \cdots\right]_j^n = 0$$

即

$$\left[\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = \frac{ah - ka^2}{2}\cdot\frac{\partial^2 u}{\partial x^2} + \left(\frac{-ah^2}{6} - \frac{a^3k^2}{3} + \frac{a^2kh}{2}\right)\cdot\frac{\partial^3 u}{\partial x^3} + \cdots\right]_j^n \tag{2.45}$$

等号右侧即为修正型方程的截断误差：

$$T_{RB} = \frac{ah}{2}(1 - a\alpha)\cdot\frac{\partial^2 u}{\partial x^2} + ah^2\left(\frac{a\alpha}{2} - \frac{1}{6} - \frac{a^2\alpha^2}{3}\right)\cdot\frac{\partial^3 u}{\partial x^3} + \cdots \tag{2.46}$$

上面推得的是迎风格式的修正型方程。显然，不同的差分格式，修正型方程的 T_R 是不一样的；同一差分格式，微分方程不同，修正型方程也不一样。从修正型方程可得出如下结论：

(1) 对同一微分方程，采用不同的差分格式，可由修正型方程的截断误差 T_R 来反映。

表 2.1

	$\frac{\partial u}{\partial t}$	$\frac{\partial u}{\partial x}$	$\frac{\partial^2 u}{\partial t^2}$	$\frac{\partial^2 u}{\partial t\partial x}$	$\frac{\partial^2 u}{\partial x^2}$	$\frac{\partial^3 u}{\partial t^3}$	$\frac{\partial^3 u}{\partial t^2\partial x}$	$\frac{\partial^3 u}{\partial t\partial x^2}$	$\frac{\partial^3 u}{\partial x^3}$	
原方程(2.44)式(1)$=\sum$右列	1	a	$\frac{k}{2}$	0	$-\frac{ah}{2}$	$\frac{k^2}{6}$	0	0	$\frac{ah^2}{6}$	$=0$
$-\frac{k}{2}\frac{\partial(1)}{\partial t}+(1)=(2)$,消去$\frac{\partial^2 u}{\partial t^2}$	1	a	0	$\frac{-ak}{2}$	$-\frac{ah}{2}$	$\frac{k^2}{6}-\frac{k^2}{4}$	0	$\frac{akh}{4}$	$\frac{ah^2}{6}$	$=0$
$\frac{k}{2}a\frac{\partial(1)}{\partial x}+(2)=(3)$,消去$\frac{\partial^2 u}{\partial t\partial x}$	1	a	0	0	$\frac{ka^2-ah}{2}$	$\frac{k^2}{6}-\frac{k^2}{4}$	$\frac{k^2 a}{4}$	$\frac{akh}{4}$	$\frac{ah^2}{6}-\frac{a^2kh}{4}$	$=0$
$\frac{k}{12}\frac{\partial^2(1)}{\partial t^2}+(3)=(4)$消去$\frac{\partial^3 u}{\partial t^3}$	1	a	0	0	$\frac{ka^2-ah}{2}$	0	$\frac{k^2 a}{12}+\frac{k^2 a}{4}$	$\frac{akh}{4}$	$\frac{ah^2}{6}-\frac{a^2kh}{4}$	$=0$
$-\frac{ak^2}{3}\frac{\partial^2(1)}{\partial t\partial x}+(4)=(5)$,消去$\frac{\partial^3 u}{\partial t^2\partial x}$	1	a	0	0	$\frac{ka^2-ah}{2}$	0	0	$\frac{akh}{4}-\frac{a^2k^2}{3}$	$\frac{ah^2}{6}-\frac{a^2kh}{4}$	$=0$
$\left(\frac{2ak^2}{3h}-\frac{k}{2}\right)\frac{\partial(1)}{\partial t}+(5)=(6)$,消去$\frac{\partial^3 u}{\partial t\partial x^2}$	1	a	0	0	$\frac{ka^2-ah}{2}$	0	0	0	$\frac{ah^2}{6}+\frac{a^3k^2}{3}-\frac{a^2kh}{2}$	$=0$

例如对于$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0$,采用 FTCS,FTFS 等离散所得修正方程的截断误差是不同的:

$$\text{FTCS}: T_{RC} = \frac{-a^2k}{2}\cdot\frac{\partial^2 u}{\partial x^2} + ah^2\left(\frac{a\alpha}{2}-\frac{1}{6}-\frac{a^2\alpha^2}{3}\right)\cdot\frac{\partial^3 u}{\partial x^3}+\cdots \tag{2.47}$$

$$\text{FTFS}: T_{RF} = \frac{-ah}{2}(1+a\alpha)\cdot\frac{\partial^2 u}{\partial x^2} + ah^2\left(\frac{a\alpha}{2}-\frac{1}{6}-\frac{a^2\alpha^2}{3}\right)\cdot\frac{\partial^3 u}{\partial x^3}+\cdots \tag{2.48}$$

$$\text{Lax-Wendroff}: T_{RL} = \frac{-h^3}{k}\cdot\frac{a\alpha-a^3\alpha^3}{6}\cdot\frac{\partial^3 u}{\partial x^3}+\frac{a^2\alpha^2h^4}{8k}(a^2\alpha^2-1)\cdot\frac{\partial^4 u}{\partial x^4}$$

$$+\frac{h^5}{k}\cdot\left(\frac{-a\alpha}{120}-\frac{a^3\alpha^3}{24}+\frac{a^5\alpha^5}{20}\right)\cdot\frac{\partial^5 u}{\partial x^5}+\cdots \tag{2.49}$$

(2) 用迎风格式解方程(1.58),当 $\sigma=a\alpha=1$ 时,$T_{RB}=0$,差分方程的解就是微分方程的解。

(3) 从 T_R 的首项可判定格式的空间精度。如迎风的空间精度为一阶精度,因为它的第一项是 $O(k,h)$;中心差分与 Lax-Wendroff 格式的空间精度为二阶精度。这样确定的精度较易理解。

2.6.2 数值耗散(Dissipation)和数值弥散(Dispersion)

为了从机理上弄清 $T_R=\sum_m D_m\frac{\partial^m u}{\partial x^m}$ 的存在对差分方程解产生的影响,我们通过对解的讨论来分析。因为 $h=\Delta x, k=\Delta t$ 都是小量,故 T_R 的后面各项均是高阶小量,主要分析其前两项,即对奇阶导数项与偶阶导数项的首项进行研究。

微分方程 $\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0$ 的解为:

$$u_1=\sum_\beta A_\beta e^{i\beta(x-at)} \tag{2.50}$$

这是微分方程的解析解,由若干个分立波迭加而成。

一、当方程右侧出现二阶或四阶空间导数项时

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=D_2\frac{\partial^2 u}{\partial x^2}+D_4\frac{\partial^4 u}{\partial x^4} \tag{2.51}$$

此方程的解为

$$u_2=\sum_\beta A_\beta e^{-D_2\beta^2t+D_4\beta^4t}e^{i\beta(x-at)} \tag{2.52}$$

即将解(2.52)代入(2.51)自然满足。分析 u_1, u_2 可以看到,A_β 是波的振幅,a 是波的传播速度;方程右侧出现的偶阶导数项没有改变波的速度和相位,只改变了波的振幅,这可由(2.52)u_2 的振幅 $A_{\beta 2}$ 得知:

$$A_{\beta 2}=A_\beta e^{-D_2\beta^2t}e^{D_4\beta^4t} \tag{2.53}$$

$$u_2=\sum_\beta A_{\beta 2}e^{i\beta(x-at)} \tag{2.54}$$

振幅 $A_{\beta 2}$ 随时间 t 的增长而变化。为保持不同 β 时 u_2 的振幅有界,必要求 $D_2>0, D_4<0$,从而使振幅在 t 增加时衰减不发散。这种由于数值计算中存在偶阶导数项截断误差而使解的模衰减的现象称为数值耗散。

二、当方程右侧出现奇阶导数项时

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = D_3\frac{\partial^3 u}{\partial x^3} + D_5\frac{\partial^5 u}{\partial x^5} \tag{2.55}$$

此方程的解为：

$$u_3 = \sum_\beta A_\beta e^{i\beta}[x - (a + D_3\beta^2 + D_5\beta^4)t] \tag{2.56}$$

可见波的振幅与(2.50)一样没受影响，但波速由 a 变为

$$a_3 = a + D_3\beta^2 + D_5\beta^4$$

波速的变化使许多分立波迭加后出现相位差异，称为数值弥散或频散。

总结上面的分析可得出如下结论：

(1) 将微分方程离散为差分方程后，通过修正方程的建立，得出的差分方程较微分方程在等号右侧增加了截断误差项。这是由离散引入的。由(2.54)式可见，其偶阶导数项使解的振幅发生耗散，为保证解的稳定，振幅只能衰减，因此，二阶导数项的前系数必须为正，四阶导数项的系数必须为负，即从振幅不增长的角度来看，要求 D_{2m} 的正负由下式决定：

$$D_{2m} = (-1)^{m+1}|D_{2m}|,\ m = 1,2,\cdots \tag{2.57}$$

差分方程的截断误差中 m 最小时首项的系数 D_{2m} 是否满足上式是判别此差分格式是否稳定的标志。例如：

T_{RB}中，$1-a\alpha>0$，$a\alpha<1$ 为稳定性条件(由(2.46)式)；

T_{RC}中，$D_2=(-a^2k/2)<0$，故无条件不稳定(由(2.47)式)；

T_{RL}中，$(a^2\alpha^2-1)<0$，$a\alpha<1$ 为稳定性条件(由(2.49)式)。

上述这种方法称为 Hirt 判稳法。它为我们在数值计算中加入人工粘性提供了依据，即在差分方程右侧加入人工粘性时，其系数的正负号应遵循(2.57)式。

(2) 求解微分方程时，解出现耗散有两种。一是微分方程本身有耗散项存在，如 $\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = \bar{\gamma}\frac{\partial^2 u}{\partial x^2}$ 中的 $\bar{\gamma}\frac{\partial^2 u}{\partial x^2}$ 项，这是物理问题本身存在的，像 N. S. 方程中的粘性耗散项，称物理耗散；二是由于数值离散时，截断误差中存在的偶阶导数项的影响，称为数值耗散，它是非物理性耗散。

又从(2.53)的分析可见，n 增加时，$t=nk$ 增加，波幅的数值耗散加剧，即衰减加剧；β 增加衰减也加快。其中 $\beta=2\pi/\lambda$ 称波数，λ 为波长，故短波衰减快，长波衰减慢。

(3) 修正方程截断误差中的奇阶导数项是弥散项，它使分立波的波速发生不同变化，出现相位差。长波弥散小，短波弥散大。奇阶导数系数 D_3、D_5 等的正负号及大小由不同格式来确定。在求解过程中，弥散可能导致解振荡，但有时振荡并不导致解的发散。

(4) 不同的差分格式实质上就是要导出修正方程中不同的奇偶阶导数项。如何使这些项能较好地反映解的真实性将决定差分格式的优劣。

2.6.3 Von Neumann 判稳的放大因子 G 与耗散和弥散之间的关系

一、一般叙述

对微分方程 $\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0$ 的精确解，我们用 Fourier 表达法写出：

$$u(j\triangle x,n\triangle t)=\sum_{\beta}A_{\beta}e^{i\beta(j\triangle x-an\triangle t)}=\sum_{\beta}A_{\beta}e^{i\beta j\triangle x}e^{-i\beta an\triangle t}$$

取一个分量，即取某一定值的 β 进行分析：

$$u(x,t)\sim A_{\beta}e^{i\beta x}e^{-i\beta ank}$$

当时间前进了 $\triangle t$ 时，

$$u(x,(n+1)k)=u(x,t+\triangle t)\sim A_{\beta}e^{i\beta x}e^{-i\beta a(n+1)k}$$

所以，对精确解来说，放大因子

$$G_e=\frac{u(x,(n+1)k)}{u(x,nk)}=e^{-i\beta ak}=|G_e|e^{i\Phi_e}$$

$$|G_e|=1,\ \Phi_e=-a\beta k=\arctan\frac{\mathrm{Im}(G_e)}{\mathrm{Re}(G_e)}\tag{2.58}$$

模为 1 即振幅没有增长；幅角 Φ_e 只表示增长 $\triangle t$ 后波行程的变化对相位的影响。

考虑差分格式 FTBS：

$$u_j^{n+1}=(1-a\alpha)u_j^n+a\alpha u_{j-1}^n$$

$$G_F=(1-a\alpha+a\alpha\cos\beta h)-ia\alpha\sin\beta h=|G_F|e^{i\Phi_F}$$

其中

$$|G_F|=[1-4a\alpha(1-a\alpha)\sin^2\frac{\beta h}{2}]^{\frac{1}{2}}\tag{2.59a}$$

$$\Phi_F=\arctan\frac{-a\alpha\sin\beta h}{1-a\alpha+a\alpha\cos\beta h}=\arg(G_F)\tag{2.59b}$$

很显然，当 $a\alpha=1$ 时，$|G_F|=1,\Phi=-\beta h=-\beta ka$，与(2.58)相比较可知，差分计算结果的放大因子与微分方程解的放大因子一致，此时解没有耗散，也没有弥散；因为 $a\alpha<1$，$\frac{|G_F|}{|G_e|}=\left[1-4a\alpha(1-a\alpha)\sin^2\frac{\beta h}{2}\right]^{\frac{1}{2}}<1$，说明在差分计算中波的振幅有不同程度的下降。正如前面所述，这是修正方程中出现偶阶导数项所引起的耗散，在 $a\alpha<1$ 的稳定性限制下，振幅不会增长。由此可见，修正方程的偶阶导数项与放大因子的模直接有关。又微分方程解的波速为 $a_e=-\frac{\Phi_e}{\beta k}$，差分方程解的波速为 $a_F=-\frac{\Phi_F}{\beta k}=\frac{\arctan\frac{-a\alpha\sin\beta h}{1-a\alpha+a\alpha\cos\beta h}}{-\beta k}$。这个波速的变化是由于修正方程中截断误差的奇阶导数项所引起的。相位误差 $\varepsilon_p=(a_F-a_e)t=a_e(q_F-1)t$，$q_F=a_F/a_e$；如 $q_F>0$ 称超前相位误差，$q_F<1$ 称滞后相位误差。

现将 Lax-Wendroff 格式的 $q_F-x(J)$ 变化关系示于图 2.6 中。J 是某解区内空间网格节点数。按定义，波长 $\lambda=2\pi/\beta_j$。在一个波长的间隔内(对周期性边界)划分网格点越多，J 越大。对周期性边界条件，在一个波长间距内，有 $\lambda=Jh=\frac{2\pi}{\beta}$，所以 $\beta h=\frac{2\pi}{J}$，J 越大，βh 越小。由(2.59)式可知，当 βh 很小时，$|G_F|$ 趋于 1，Φ_F 趋于 $-a\alpha\beta h$，于是 q_F 也趋于 1。当 βh

增加时，q_F 越来越小，且趋于 0；对超前相位误差的格式，情况正好相反。从图 2.6 可以看出，Lax-Wendroff格式是滞后相位误差。

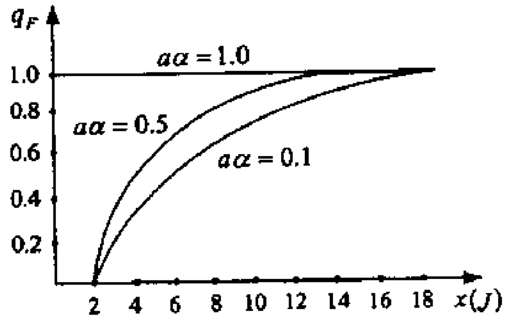


图 2.6

二、由 G 导出修正方程截断误差

对于微分方程：

$$\frac{\partial u}{\partial t}+L_x(u)=0 \qquad (2.60)$$

差分方程为

$$\left\{\frac{\partial u}{\partial t}+L_x(u)\right\}_j^n=\sum_m D_m\left\{\frac{\partial^m u}{\partial x^m}\right\}_j^n \qquad (2.61a)$$

或

$$\frac{\partial u}{\partial t}=\sum_{m=0}^{\infty}D_{2m+1}\frac{\partial^{2m+1}u}{\partial x^{2m+1}}+\sum_{m=1}^{\infty}D_{2m}\frac{\partial^{2m}u}{\partial x^{2m}} \qquad (2.61b)$$

差分方程的解为

$$u_j^n=u(nk,jh)\sim A_\beta e^{i\beta(x-c_0t)}=e^{\alpha t}e^{i\beta x} \qquad (2.62)$$

其中 $e^{\alpha t}=A_\beta e^{-i\beta c_0 t}$，$\alpha=r+i\delta$ 是复数。

将 $u_j^n=e^{\alpha t}e^{i\beta x}$代入(2.61b)得：

$$\alpha=i\sum_{m=0}^{\infty}D_{2m+1}(-1)^m\beta^{2m+1}+\sum_{m=1}^{\infty}D_{2m}(-1)^m\beta^{2m} \qquad (2.63)$$

故：

$$r=\sum_{m=1}^{\infty}D_{2m}(-1)^m\beta^{2m},\ \delta=\sum_{m=0}^{\infty}D_{2m+1}(-1)^m\beta^{2m+1} \qquad (2.64)$$

又

$$G_F=|G_F|e^{i\Phi_F}=u_j^{n+1}/u_j^n=e^{\alpha k}=e^{(r+i\delta)k}=e^{rk}e^{i\delta k}$$

所以，

$$\ln|G_F|=rk,\ e^{i\Phi_F}=e^{i\delta k}$$

由(2.64)得：

$$r=\frac{\ln|G_F|}{k}=\sum_{m=1}^{\infty}D_{2m}(-1)^m\beta^{2m} \qquad (2.65a)$$

$$\delta=\frac{\Phi_F}{k}=\sum_{m=0}^{\infty}D_{2m+1}(-1)^m\beta^{2m+1} \qquad (2.65b)$$

(2.65)式中的 $\ln|G_F|$ 与 Φ_F 均可展开成多项级数，即：

$$\ln|G_F|=\frac{1}{2}\ln|G_F|^2=\frac{1}{2}\ln[1-(1-|G_F|^2)]=\sum_{m=1}^{\infty}E_{2m}\beta^{2m}h^{2m} \qquad (2.66a)$$

$$\Phi_F=\arctan\frac{\mathrm{Im}(G_F)}{\mathrm{Re}(G_F)}=\sum_{m=0}^{\infty}F_{2m+1}\beta^{2m+1}h^{2m+1} \qquad (2.66b)$$

比较(2.66)与(2.65)式两侧的系数，即可求出 D_{2m}，D_{2m+1}。从而求出截断误差。下面以方程(1.58)的 FTBS 差分方程为例，导出修正方程，得到截断误差的前两项。由(1.90)

得：$\ln|G_F| = \frac{1}{2}\ln\{1-[4\sigma(1-\sigma)\sin^2\frac{\beta h}{2}]\}$，展开后可得 E_{2m}值，从而由(2.65a)和(2.66a)可求得：

$$D_{2m} = \frac{(-1)^m}{k}[E_{2m}h^{2m}], \quad m = 1,2,\cdots$$

同理，由(1.90)得：$\Phi_F = \arctan\frac{-\sigma\sin\beta h}{1-\sigma+\sigma\cos\beta h}$，展开后可得 F_{2m+1}值，从而由(2.65b)和(2.66b)可求得：

$$D_{2m+1} = \frac{(-1)^m}{k}[F_{2m+1}h^{2m+1}], \quad m = 0,1,2,\cdots$$

由 D_{2m}，D_{2m+1}即可构成截断误差。我们感兴趣的是截断误差的首两项。

2.7 Burgers 方程的数值解法

考虑 N.S. 方程的模型方程(2.26)式，即：$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \bar{\gamma}\frac{\partial^2 u}{\partial x^2}$，其线化后的形式为：

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = b\frac{\partial^2 u}{\partial x^2}, \quad a,b\ 均为常数 \tag{2.67}$$

(2.67)是线性方程，便于问题的分析。

2.7.1 差分格式的稳定性和网格雷诺数

如果对(2.67)用时间空间均为一阶精度的差分格式，那么半离散修正型方程(指时间导数不离散且为零，只有空间导数离散的修正方程)的截断误差首项是 $D_2\frac{\partial^2 u}{\partial x^2}$，这是数值耗散项。如果 b 不太大就易使数值耗散掩盖物理耗散，使解不合理。因此，常采用具有二阶精度的差分格式求解(2.67)。下面选取 FTCS 进行分析：

$$\frac{1}{k}(u_j^{n+1} - u_j^n) + \frac{a}{2h}(u_{j+1}^n - u_{j-1}^n) = \frac{b}{h^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \tag{2.68}$$

差分方程的相容性容易证明。下面用 Von Neumann 法分析其稳定性。令 $r=bk/h^2$，$\sigma=ak/h$ 则：

$$u_j^{n+1} = (r+\frac{\sigma}{2})u_{j-1}^n + (1-2r)u_j^n + (r-\frac{\sigma}{2})u_{j+1}^n$$

$$G = (r+\frac{\sigma}{2})e^{-i\beta h} + (1-2r) + (r-\frac{\sigma}{2})e^{i\beta h}$$

$$= 2r\cos\beta h + (1-2r) - i\sigma\sin\beta h$$

令 $G=G_R+iG_I$，则$\left[\frac{G_R-(1-2r)}{2r}\right]^2+\left[\frac{G_I}{\sigma}\right]^2=1$。这是一个以 $2r$ 为长轴，σ 为短轴的椭圆，其中心在$(1-2r,0)$处。为保证稳定，它必须被包在$|G|=1$的圆内，因此，稳定的充分条件是 $2r\leqslant1$，$\sigma\leqslant1$，$\sigma\leqslant2r$，即 $\sigma\leqslant2r\leqslant1$。

由 $\sigma\leqslant1$ 得：$k\leqslant\frac{h}{a}$

由 $2r\leqslant1$ 得：$k\leqslant\frac{h^2}{2b}$

所以，其稳定性条件为：

$$k \leqslant \min\left\{\frac{h}{a}, \frac{h^2}{2b}\right\} \tag{2.69}$$

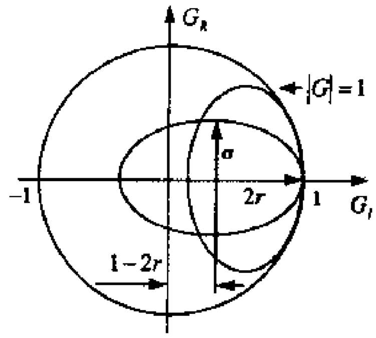


图 2.7

从作图的情况看，$\sigma \leqslant 2r$ 并非必要条件。但数值计算表明，$\sigma \leqslant 2r$ 意味着：$\frac{ka}{h} \leqslant \frac{2kb}{h^2}$，即$\frac{ah}{b} \leqslant 2$。比较(2.67)与(2.26)得$\frac{ah}{b} = \frac{ah}{\bar{\gamma}} \leqslant 2$。由(2.26)知，$\bar{\gamma} = \frac{\mu}{\rho}$，所以有：

$$\frac{\rho ah}{\mu} \leqslant 2 \tag{2.70}$$

其中 a 表示速度，h 表示网格长度，因此，由(2.70)知，$\frac{ah}{b}$是以网格间距 h 为特征长度的雷诺数，称为网格雷诺数，记为 $\mathrm{Re}(\triangle x)=\frac{ah}{b}$。$\frac{ah}{b} \leqslant 2$ 理解为网格雷诺数小于等于 2。因此，$\sigma \leqslant 2r$，或 $\mathrm{Re}(\triangle x) \leqslant 2$ 反映了流动的物理状况与数值计算中网格的相互关系。$\mathrm{Re}(\triangle x)$的大小不影响稳定性，但下面的分析表明，$\mathrm{Re}(\triangle x)>2$ 时，在物理量有突跃处(如激波前后)，数值解将出现振荡。这是格式本身的弥散引起的，与稳定性无关。下节进行详细分析。

2.7.2 网格雷诺数对数值振荡的影响

以(2.67)为例，并考虑定常状态：

$$a\frac{\mathrm{d}u}{\mathrm{d}x} - b\frac{\mathrm{d}^2u}{\mathrm{d}x^2} = 0,\ b>0,\ 0 \leqslant x \leqslant 1 \tag{2.71}$$

边界条件：$u|_{x=0} = u_0,\ u|_{x=1} = u_1$

(2.71)的精确解为：

$$u = u_0 + (u_1 - u_0)(e^{\frac{ax}{b}} - 1)/(e^{\frac{a}{b}} - 1) \tag{2.72}$$

对(2.71)采用中心差分得：

$$\frac{a}{2\triangle x}(u_{j+1} - u_{j-1}) - \frac{b}{\triangle x^2}(u_{j+1} - 2u_j + u_{j-1}) = 0$$

即：

$$\left(2 - \frac{a\triangle x}{b}\right)u_{j+1} - 4u_j + \left(2 + \frac{a\triangle x}{b}\right)u_{j-1} = 0 \tag{2.73}$$

分析(2.73)的特解。令 $u_j = R^j$，$\frac{a\triangle x}{b} = \mathrm{Re}(\triangle x)$ 则有：

$$[2 - \mathrm{Re}(\triangle x)]R^2 - 4R + [2 + \mathrm{Re}(\triangle x)] = 0$$

求解上式得：

$$R = 1,\ R = \frac{2 + \mathrm{Re}(\triangle x)}{2 - \mathrm{Re}(\triangle x)}$$

所以，(2.73)的通解为：

$$u_j = A + B\left[\frac{2 + \mathrm{Re}(\triangle x)}{2 - \mathrm{Re}(\triangle x)}\right]^j \tag{2.74}$$

下面通过比较差分方程的解(2.74)和微分方程的解(2.72)进行分析。

一、$\mathrm{Re}(\triangle x)\ll 1$

(2.74)式可写成：$u_j=A+Be^{j\ln\left[\frac{2+\mathrm{Re}(\triangle x)}{2-\mathrm{Re}(\triangle x)}\right]}$。用级数展开，得：

$$u_j = A + Be^{j\mathrm{Re}(\triangle x)} = A + Be^{j\frac{a\triangle x}{b}} \tag{2.75}$$

将边界条件代入(2.75)可得：

$$u = u_0 + (u_1 - u_0)(e^{\frac{ax}{b}} - 1)/(e^{\frac{a}{b}} - 1)$$

此即(2.72)。所以，当 $\mathrm{Re}(\triangle x)\ll 1$ 时，差分方程的解即为微分方程的精确解。

二、$\mathrm{Re}(\triangle x)\leqslant 2$

为讨论方便，令边界条件 $u_0=0, u_1=-1$ 则微分方程的解为：

$$u = -\frac{\left(e^{\frac{ax}{b}} - 1\right)}{\left(e^{\frac{a}{b}} - 1\right)}$$

令(2.74)中的 $\dfrac{2+\mathrm{Re}(\triangle x)}{2-\mathrm{Re}(\triangle x)} = R$，且当 $j=0$ 时，$u_0=0$，$j=J$ 时，$u_J=-1$。将 u_0, u_J 代入(2.74)得：$A=-1/(1-R^J)$，$B=1/(1-R^J)$。将 A,B 代入(2.74)得：

$$u_j = -\frac{\left\{1-\left[\dfrac{2+\mathrm{Re}(\triangle x)}{2-\mathrm{Re}(\triangle x)}\right]^j\right\}}{\left\{1-\left[\dfrac{2+\mathrm{Re}(\triangle x)}{2-\mathrm{Re}(\triangle x)}\right]^J\right\}} \tag{2.76}$$

当 $\mathrm{Re}(\triangle x)\leqslant 2$ 时，(2.76)式右侧为小于零的值，由于 $j<J$，所以 u 由 0 光滑地变化到 -1（见图 2.8），差分解单调收敛于微分方程的解。

三、$\mathrm{Re}(\triangle x)>2$

微分方程的解与差分方程的解的公式同(2.76)和(2.72)式，但差分方程(2.76)的解中分子式 $\{[2+\mathrm{Re}(\triangle x)]/[2-\mathrm{Re}(\triangle x)]\}^j$ 的值由于当 j 为偶数时为正，j 为奇数时为负，所以差分方程的解随着 j 的变化而振荡。此振荡随 $\mathrm{Re}(\triangle x)$ 的增大而加剧，且这种振荡不是由稳定性引起的，而是由数值弥散引起的。

上述解析解与差分解示于图 2.8 中。

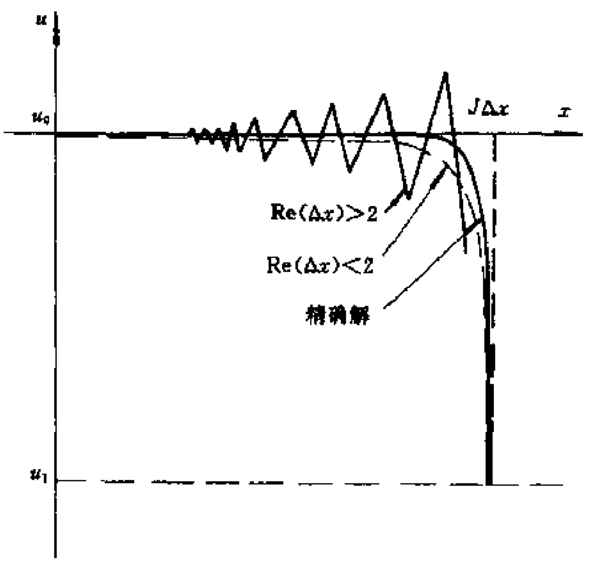


图 2.8

实际流中，$\mathrm{Re}\approx 10^6$，$h\approx \frac{1}{100}\sim\frac{1}{500}$，$\mathrm{Re}(\triangle x)\gg 2$，由弥散引起的数值振荡是不可避免的。如用迎风格式，则用二阶数值耗散去控制三阶弥散，消去振荡，但物理耗散又受到了数值耗散过大的影响。于是有人提出用高

于二阶空间精度的格式求解N. S. 方程。

2.7.3 数值弥散与人工粘性

为了进一步说明 Re(△x) 大于 2 时，数值弥散引起的振荡，可由 FTCS 差分格式离散 $\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=\bar{\gamma}\frac{\partial^2 u}{\partial x^2}$ 后的差分方程进行分析。差分方程是：

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}-\bar{\gamma}\frac{\partial^2 u}{\partial x^2}+D_3\frac{\partial^3 u}{\partial x^3}+D_4\frac{\partial^4 u}{\partial x^4}+\cdots=0 \tag{2.77}$$

设定常状态下，且使 $a=1$ 后略去高阶项，得：

$$\frac{\mathrm{d}u}{\mathrm{d}x}-\bar{\gamma}\frac{\mathrm{d}^2 u}{\mathrm{d}x^2}+D_3\frac{\mathrm{d}^3 u}{\mathrm{d}x^3}=0 \tag{2.78}$$

特征方程：

$$\lambda-\bar{\gamma}\lambda^2+D_3\lambda^3=0$$

$$\lambda_1=0,\ \lambda_{2,3}=\frac{\bar{\gamma}\pm\sqrt{\bar{\gamma}^2-4D_3}}{2D_3} \tag{2.79}$$

下面对(2.78)的解 u 进行分析。

(1) 若 $\bar{\gamma}^2>4D_3$，λ_2，λ_3 为实数，$u=C_1e^{\lambda_2 x}+C_2e^{\lambda_3 x}+C_3$，代入边界条件后得到光滑的解。

(2) 若 $\bar{\gamma}^2<4D_3$，$\lambda_2=A+Bi$，$\lambda_3=A-Bi$，$A=\frac{\bar{\gamma}}{2D_3}$，$B=\frac{\sqrt{4D_3-\bar{\gamma}^2}}{2D_3}$，$u=e^{Ax}\cos Bx+Ce^{Ax}\sin Bx$，出现振荡解。因此，只有当 $\bar{\gamma}^2>4D_3$，$\bar{\gamma}>2\sqrt{D_3}$，即二阶耗散项系数按规律大于三阶弥散项时，耗散才能阻尼弥散，得到光滑解。可以证明，$\bar{\gamma}>2\sqrt{D_3}$ 与 Re(△x)<2 的要求是相当的。

本节小结：

(1) 对 Burgers 方程差分计算的分析表明，如对流项采用二阶中心差分，则当 Re(△x)>2 时，即定常情况下假设 $a=1$ 时，$\bar{\gamma}^2<4D_3$，由于二阶耗散不足以阻尼三阶弥散，在物理量有较大变化处解发生振荡。

(2) 如果对流项采用一阶迎风格式，则二阶数值耗散增加，足以阻尼三阶弥散，解不会出现振荡，但数值耗散掩盖了物理耗散，可能使解失去意义 。

(3) a 如不等于 1，则 $\lambda_{2,3}=\frac{\bar{\gamma}\pm\sqrt{\bar{\gamma}^2-4aD_3}}{2D_3}$，因此 a 的大小和符号也会对振荡有一定影响，但上面获得的机理性分析具有一般意义。

2.8 守恒型双曲型方程组的求解

2.8.1 一维守恒型双曲型方程组(Euler 方程组)

由(1.1)式略去 y 向导数与粘性项得：

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = 0, \quad -\infty < x < \infty \tag{2.80}$$

初值问题：

$$U_0 = f(x,0)$$

其中

$$U = (\rho, \rho u, \rho E)^T,$$

$$F = F_I = [\rho u, \rho u^2 + p, (\rho E + p)u]^T,$$

$$E = \frac{p}{(\gamma - 1)\rho} + \frac{u^2}{2} = \frac{E_1}{\rho}$$

一维双曲型方程组求解时为什么要保持守恒形式呢？至少有下列几个原因：

(1) 初值条件下，守恒型方程可唯一求得间断解。

大家知道，在超声速流动中，定常二维或一维流中均有激波存在，一维非定常流中也有激波存在。只有在守恒的微分方程与守恒的差分方程中才能较好地求得激波。为此，我们先从物理概念上简单地加以说明。

设气流跨过一个定常正激波，见图 2.9，通常以单个原始变量 ρ, p 等来分析跨激波前后的变化。由图(a)可见，激波后的 ρ_2 比激波前的 ρ_1 大，在激波处有突跃；同样激波后的 p_2 比激波前的 p_1 大，也在激波处有突跃(见图(c))。

非守恒方程以原参数 ρ, p, u, v 为主变量，在激波处出现不连续突跃，于是在数值计算中会引起一系列数值误差，从而影响激波的正确求解。

用守恒形式的方程求解，从图(b)与(d)可见，按质量守恒，激波前后：

$$\rho_1 u_1 = \rho_2 u_2$$

按动量守恒，激波前后：

$$p_1 + \rho_1 u_1^2 = p_2 + \rho_2 u_2^2$$

这些组合参数在激波处是连续的，无突跃。由此可见用守恒变量进行数值计算，在捕捉激波时误差较小，稳定性较好。在守恒型控制方程中，变量通量就是上述组合变量，在跨激波的流动中变动甚微。在超声速流与高超声速流动中激波总是存在的，用守恒型的控制方程与守恒的差分格式尤其重要。

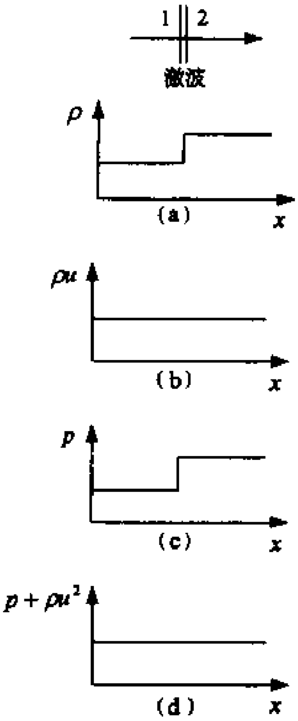


图 2.9

下面用线积分的方法，从数学上叙述一下：方程(2.80)在初值条件下可用多种途径唯一地求得弱解。

如图 2.10 所示，在 $t>0$ 平面上取任意域 R，边界 ∂R，则由(2.80)得：

$$\iint_R \left(\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} \right) \mathrm{d}x\mathrm{d}t = 0 \tag{2.81}$$

由 Green 公式得：

$$\int_{\partial R} (U\mathrm{d}x - F\mathrm{d}t) = 0 \tag{2.82}$$

方程(2.80)的广义解是满足初值$U_0=f(x,0)$，且在∂R上满足(2.82)的任何$U(x,t)$函数，这种解称为弱解，$U(x,t)$在R上并不要求处处连续可微，允许有突变。激波就是一个例子。现证明之。

设R内任意曲线C上变量$U(x,t)$不连续，现导出C前后U的变化规律：

$$\int_{\partial R}(U\mathrm{d}x-F\mathrm{d}t)=\int_{\partial R_1}+\int_{C_1}+\int_{C_2}+\int_{\partial R_2}=0$$

所以

$$\int_{C_1}(U\mathrm{d}x-F\mathrm{d}t)+\int_{C_2}(U\mathrm{d}x-F\mathrm{d}t)=0$$

$$\oint_C(U\mathrm{d}x-F\mathrm{d}t)=0$$

$$\oint_C[(U_2-U_1)\mathrm{d}x-(F_2-F_1)\mathrm{d}t]=0 \tag{2.83}$$

$$\oint_C([U]\mathrm{d}x-[F]\mathrm{d}t)=0$$

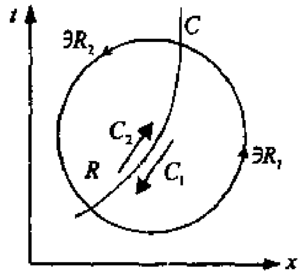


图 2.10

$[U]$，$[F]$表示跨物理量不连续线C的物理量的突跃值。上述结论(2.83)在R域上任一处均合理。因为$\frac{\mathrm{d}x}{\mathrm{d}t}=v_s$是波速，故由(2.83)$v_s=\frac{[F]}{[U]}$，从而得：

$$\begin{aligned}v_s\cdot[\rho]&=[\rho u]\\ v_s\cdot[\rho u]&=[p+\rho u^2]\\ v_s\cdot[\rho E]&=[(\rho E+p)u]\end{aligned} \tag{2.84}$$

(2.84)式称为激波关系式。以(2.84)第一式为例：

$$(\rho_2-\rho_1)v_s=\rho_2u_2-\rho_1u_1$$

在静止气体中，激波以v_s前进，激波前后$u_1=0$，$u_2=v_g$，v_g是气体受扰动后的前进速度。所以，

$$v_s=\frac{\rho_2v_g}{\rho_2-\rho_1}$$

这个公式与气体动力学的公式是一致的。但是，如果将质量方程写成非守恒形式$\frac{\partial\rho}{\partial t}+\rho\frac{\partial u}{\partial x}+u\frac{\partial\rho}{\partial x}=0$，就不能正确地解出弱解。

(2) 守恒型方程的离散保证计算区域内不会有误差积累。

以二维质量守恒方程为例，解区是一个矩形域，如图2.11所示。积分形式方程为：

$$\int_S\frac{\partial\rho}{\partial t}\mathrm{d}s+\int_{\partial S}\rho v_n\mathrm{d}l=0$$

微分形式方程：

$$\frac{\partial\rho}{\partial t}+\frac{\partial\rho v_x}{\partial x}+\frac{\partial\rho v_y}{\partial y}=0$$

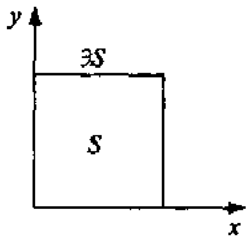


图 2.11

从物理意义上说，控制体内密度的时间变化率等于控制体边界∂S上质量的净流出。在控制体内没有质量流的堆积和出现真空。先按守恒形式写出：

$$\frac{\rho_{1,1}^{n+1}-\rho_{1,1}^{n}}{\triangle t}=-\left[\frac{\rho_{1,1}^{n}u_{1,1}^{n}-\rho_{0,1}^{n}u_{0,1}^{n}}{\triangle x}+\frac{\rho_{1,1}^{n}v_{1,1}^{n}-\rho_{1,0}^{n}v_{1,0}^{n}}{\triangle y}\right]$$

$$\frac{\rho_{1,2}^{n+1}-\rho_{1,2}^{n}}{\triangle t}=-\left[\frac{\rho_{1,2}^{n}u_{1,2}^{n}-\rho_{0,2}^{n}u_{0,2}^{n}}{\triangle x}+\frac{\rho_{1,2}^{n}v_{1,2}^{n}-\rho_{1,1}^{n}v_{1,1}^{n}}{\triangle y}\right]$$

$$\frac{\rho_{2,1}^{n+1}-\rho_{2,1}^{n}}{\triangle t}=-\left[\frac{\rho_{2,1}^{n}u_{2,1}^{n}-\rho_{1,1}^{n}u_{1,1}^{n}}{\triangle x}+\frac{\rho_{2,1}^{n}v_{2,1}^{n}-\rho_{2,0}^{n}v_{2,0}^{n}}{\triangle y}\right]$$

$$\vdots$$

$$\frac{\rho_{J,M}^{n+1}-\rho_{J,M}^{n}}{\triangle t}=-\left[\frac{\rho_{J,M}^{n}u_{J,M}^{n}-\rho_{J-1,M}^{n}u_{J-1,M}^{n}}{\triangle x}+\frac{\rho_{J,M}^{n}v_{J,M}^{n}-\rho_{J,M-1}^{n}v_{J,M-1}^{n}}{\triangle y}\right]$$

上述各式右侧有许多项可对应消去，从而得：

$$\triangle S\sum_{j,m}\frac{\rho_{j,m}^{n+1}-\rho_{j,m}^{n}}{\triangle t}=\left\{\sum\frac{\rho_{j,0}^{n}v_{j,0}^{n}}{\triangle y}+\sum\frac{\rho_{0,m}^{n}u_{0,m}^{n}}{\triangle x}-\sum\frac{\rho_{J,m}^{n}v_{J,m}^{n}}{\triangle y}-\sum\frac{\rho_{j,M}^{n}u_{j,M}^{n}}{\triangle x}\right\}\triangle S$$

即：

$$\int_S\frac{\partial\rho}{\partial t}\mathrm{d}s+\int_{\partial S}\rho v_n\mathrm{d}l=0$$

它与积分方程形式一致，控制体内没有误差积累。

对非守恒型方程：

$$\frac{\partial\rho}{\partial t}=-\left(\rho\frac{\partial u}{\partial x}+u\frac{\partial\rho}{\partial x}+\rho\frac{\partial v}{\partial y}+v\frac{\partial\rho}{\partial y}\right)$$

第一项：

$$\frac{\rho_{1,1}^{n+1}-\rho_{1,1}^{n}}{\triangle t}=-\left[\rho_{1,1}^{n}\frac{u_{1,1}^{n}-u_{0,1}^{n}}{\triangle x}+u_{1,1}^{n}\frac{\rho_{1,1}^{n}-\rho_{0,1}^{n}}{\triangle x}+\rho_{1,1}^{n}\frac{v_{1,1}^{n}-v_{1,0}^{n}}{\triangle y}+v_{1,1}^{n}\frac{\rho_{1,1}^{n}-\rho_{1,0}^{n}}{\triangle y}\right]$$

第二项：

$$\frac{\rho_{1,2}^{n+1}-\rho_{1,2}^{n}}{\triangle t}=-\left[\rho_{1,2}^{n}\frac{u_{1,2}^{n}-u_{0,2}^{n}}{\triangle x}+u_{1,2}^{n}\frac{\rho_{1,2}^{n}-\rho_{0,2}^{n}}{\triangle x}+\rho_{1,2}^{n}\frac{v_{1,2}^{n}-v_{1,1}^{n}}{\triangle y}+v_{1,2}^{n}\frac{\rho_{1,2}^{n}-\rho_{1,1}^{n}}{\triangle y}\right]$$

……

由此看出，各项之间不可能互相消去，最后在控制体内有误差的积累，在迭代过程中误差增大。

(3) 守恒型方程导出的守恒型差分方程可克服某些非线性不稳定性。

2.8.2 守恒型的差分格式

考虑方程组(2.81)，积分得：

$$\int_{(j-\frac{1}{2})h}^{(j+\frac{1}{2})h}(U^{n+1}-U^{n})\mathrm{d}x+\int_{nk}^{(n+1)k}(F_{j+\frac{1}{2}}-F_{j-\frac{1}{2}})\mathrm{d}t=0\qquad(2.85)$$

通常

$$\int_{(j-\frac{1}{2})h}^{(j+\frac{1}{2})h}(U^{n+1}-U^{n})\mathrm{d}x\approx(U_j^{n+1}-U_j^{n})\triangle x_j$$

$$\int_{nk}^{(n+1)k}(F_{j+\frac{1}{2}}-F_{j-\frac{1}{2}})\mathrm{d}t\approx(F_{j+\frac{1}{2}}^{n+\frac{1}{2}}-F_{j-\frac{1}{2}}^{n+\frac{1}{2}})\triangle t_n$$

其中 $F_{j+\frac{1}{2}}^{n+\frac{1}{2}}$,$F_{j-\frac{1}{2}}^{n+\frac{1}{2}}$表示单元中的某种平均,根据求解情况的变化而异。于是得守恒型差分格式:

$$U_j^{n+1} = U_j^n - \alpha(F_{j+\frac{1}{2}}^{n+\frac{1}{2}} - F_{j-\frac{1}{2}}^{n+\frac{1}{2}}),\ \alpha = \frac{\Delta t}{\Delta x} \tag{2.86}$$

一般

$$F_{j+\frac{1}{2}}^{n+\frac{1}{2}} = f(u_{j+l}^n,\ u_{j+l-1}^n, \cdots, u_{j-l+2}^n,\ u_{j-l+1}^n)$$

$$F_{j-\frac{1}{2}}^{n+\frac{1}{2}} = f(u_{j+l-1}^n, u_{j+l-2}^n, \cdots, u_{j-l+1}^n, u_{j-l}^n)$$

犹如特征线插值关系式,以 j 为基准,所取的点 $j+1/2$ 比 $j-1/2$ 右移一个网格。相容性要求:

$$f(u,u,\cdots,u) = F(u) \tag{2.87}$$

利用守恒型差分方程数值计算既可达到求出激波、减少误差积累的要求,又可保证差分方程的收敛性。

下面由(2.86)和(2.87)出发,分析一些常用的守恒型差分格式。

一、显式中心差分

令

$$F_{j+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(F_j^n + F_{j+1}^n) = \frac{1}{2}[F(U_j^n) + F(U_{j+1}^n)]$$

$$F_{j-\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(F_j^n + F_{j-1}^n) = \frac{1}{2}[F(U_j^n) + F(U_{j-1}^n)]$$

代入(2.86)得:

$$U_j^{n+1} = U_j^n - \frac{\alpha}{2}(F_{j+1}^n - F_{j-1}^n) \tag{2.88}$$

且 $F_{j\pm\frac{1}{2}}^{n+\frac{1}{2}}(U_j,U_j) = F(U_j)$,故与差分方程相容,满足(2.87)式。

(2.88)显然是中心差分格式,它是守恒的,精度为时间一阶、空间二阶,但不稳定,故不能使用。

二、耗散型中心差分显格式

$$F_{j+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(F_j^n + F_{j+1}^n) - \frac{\Delta t}{2\Delta x}(U_{j+1}^n - U_j^n)$$

$$F_{j-\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(F_j^n + F_{j-1}^n) - \frac{\Delta t}{2\Delta x}(U_j^n - U_{j-1}^n)$$

代入(2.86)得:

$$U_j^{n+1} = U_j^n - \frac{\alpha}{2}(F_{j+1}^n - F_{j-1}^n) + \frac{\alpha^2}{2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n) \tag{2.89}$$

且 $F_{j\pm\frac{1}{2}}^{n+\frac{1}{2}}(U_j,U_j)=F(U_j)$,故与差分方程相容。

(2.89)称为 Lax 的耗散格式。若令 $A=\frac{\partial F}{\partial U}$,则(2.89)变为:

$$U_j^{n+1} = U_j^n - \frac{\alpha A}{2}(U_{j+1}^n - U_{j-1}^n) + \frac{\alpha^2}{2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n)$$

此格式的目的是在(2.88)的基础上增加耗散项以达到稳定的要求。

三、Lax-Wendroff 型格式

对标量方程(1.58)可导出：

$$u_j^{n+1}=u_j^n+\left(\frac{\partial u}{\partial t}\right)_j^n k+\left(\frac{\partial^2 u}{\partial t^2}\right)_j^n\frac{k^2}{2}+O(k^3)$$

$$=u_j^n-ak\left(\frac{\partial u}{\partial x}\right)_j^n+\frac{a^2k^2}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_j^n+O(k^3)$$

对双曲型方程组(2.80)，设 $A=\dfrac{\partial F}{\partial U}$，同理有：

$$U_j^{n+1}=U_j^n-k\left(\frac{\partial F}{\partial x}\right)_j^n+\frac{k^2}{2}\frac{\partial}{\partial x}\left(A\frac{\partial F}{\partial x}\right)_j^n \tag{2.90}$$

所以，Lax-Wendroff 格式为：

$$U_j^{n+1}=U_j^n-\frac{\alpha}{2}(F_{j+1}^n-F_{j-1}^n)+\frac{\alpha^2}{2}\left[A_{j+\frac{1}{2}}^n(F_{j+1}^n-F_j^n)-A_{j-\frac{1}{2}}^n(F_j^n-F_{j-1}^n)\right] \tag{2.91}$$

$\alpha=\dfrac{k}{h}$，$A_{j\pm\frac{1}{2}}^n=A\left(\dfrac{U_{j\pm1}^n+U_j^n}{2}\right)$表示矩阵 A 内各元素由$\dfrac{U_{j\pm1}^n+U_j^n}{2}$构成。

又(2.90)可写成：

$$U_j^{n+1}=U_j^n-k\left(\frac{\partial F}{\partial x}\right)_j^n+\frac{k^2}{2}\frac{\partial}{\partial x}\left(A^2\frac{\partial U}{\partial x}\right)_j^n$$

所以，(2.91)又可写成：

$$U_j^{n+1}=U_j^n-\frac{\alpha}{2}(F_{j+1}^n-F_{j-1}^n)+\frac{\alpha^2}{2}$$

$$\left[(A_{j+\frac{1}{2}}^n)^2(U_{j+1}^n-U_j^n)-(A_{j-\frac{1}{2}}^n)^2(U_j^n-U_{j-1}^n)\right] \tag{2.92}$$

(2.91)或(2.92)是相容的守恒型差分格式，因为它可由

$$F_{j+\frac{1}{2}}^{n+\frac{1}{2}}=\frac{1}{2}(F_{j+1}^n+F_j^n)-\frac{\alpha}{2}A_{j+\frac{1}{2}}^n(F_{j+1}^n-F_j^n)$$

$$F_{j-\frac{1}{2}}^{n+\frac{1}{2}}=\frac{1}{2}(F_{j-1}^n+F_j^n)-\frac{\alpha}{2}A_{j-\frac{1}{2}}^n(F_j^n-F_{j-1}^n)$$

代入(2.86)而得到，而上式与原方程相容，格式的精度是(k^2,h^2)，稳定性条件为 $\alpha\rho(A)\leqslant 1$。

四、MacCormack 显式二步格式

在上述 Lax-Wendroff 格式计算中，A 矩阵的处理是一个麻烦的问题。如方程组是由 m 个列矢量组成，则 A 是 $m\times m$ 矩阵。而 $A_{j+\frac{1}{2}}^n$ 是由$U_{j+\frac{1}{2}}^n$ 所组成，即 $A_{j+\frac{1}{2}}^n=A(U_{j+\frac{1}{2}}^n)$，每次运算量很大，尤其对非线性方程。1969 年 MacCormack 提出二步格式：

$$U_j^{\overline{n+1}}=U_j^n-\alpha(F_{j+1}^n-F_j^n) \tag{2.93a}$$

$$U_j^{\overline{\overline{n+1}}}=U_j^{\overline{n+1}}-\alpha(F_j^{\overline{n+1}}-F_{j-1}^{\overline{n+1}}) \tag{2.93b}$$

$$U_j^{n+1}=\frac{1}{2}(U_j^n+U_j^{\overline{\overline{n+1}}}) \tag{2.93c}$$

这是由 Lax-Wendroff 格式变形而来。上面列出的是先前差再后差的 FB 格式，BF 格式也可写出。在常系数矩阵情况下，线化代入，证明这是守恒型相容的差分格式。时间与空间精度均为二阶，稳定性条件是 $\alpha\rho(A)\leqslant 1$。

五、一般二步守恒显格式

$$U_j^{\overline{n+1}} = (1-r)U_j^n + rU_{j+1}^n - \delta\alpha(F_{j+1}^n - F_j^n) \tag{2.94a}$$

$$U_j^{n+1} = U_j^n - \frac{\alpha}{2\delta}[(\delta-r)F_{j+1}^n + (2r-1)F_j^n + (1-r-\delta)F_{j-1}^n + F_j^{\overline{n+1}} - F_{j-1}^{\overline{n+1}}] \tag{2.94b}$$

式中 $F_j^{\overline{n+1}} = F(U_j^{\overline{n+1}})$，$\gamma$、$\delta$ 是在 0 到 1 间可调的常数。

(1) 当 $\delta=r=1$ 时，即为 MacCormack BF 格式；

(2) 当 $\delta=1, r=0$ 时，即为 MacCormack FB 格式；

(3) 当 $\delta=r=1/2$ 时，即为 Rithtmyer 格式：

$$U_j^{\overline{n+1}} = \frac{1}{2}(U_{j+1}^n + U_j^n) - \frac{\alpha}{2}(F_{j+1}^n - F_j^n) \tag{2.95a}$$

$$U_j^{n+1} = U_j^n - \alpha(F_j^{\overline{n+1}} - F_{j-1}^{\overline{n+1}}) \tag{2.95b}$$

4. 当 $\delta=1, r=1/2$ 时，即为 Rabin-Burstain 格式。

此外，还可以形成其他一些格式。这些格式的共同特点是：

① 显式守恒型相容格式；

② 预测、校正二步格式；

③ 三点 $(j+1, j, j-1)$ 二层 $(n, n+1)$ 格式；

④ 时间和空间的精度均为二阶。

它们的差别是修正方程截断误差中各阶导数前的系数不同。

(2.94)式的修正型方程为：

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} =$$

$$\frac{h^2}{6}\cdot\frac{\partial}{\partial x}\left\{(\alpha^2A^2-1)\frac{\partial^2 F}{\partial x^2} - \frac{3}{2\delta}\frac{dA}{dU}\frac{\partial}{\partial x}(rU-\delta\alpha F)\frac{\partial}{\partial x}[(r-1)U-\delta\alpha F] + 2\alpha^2\frac{dA}{dU}\left(\frac{\partial F}{\partial x}\right)^2\right\} \tag{2.96}$$

以方程 $\dfrac{\partial U}{\partial t} + \dfrac{\partial(U^2/2)}{\partial x} = 0$ 为例，$F = \dfrac{U^2}{2}$，由(2.96)知，修正方程的截断误差为：

$$T_R = h^2\left[E_3\frac{\partial^3 U}{\partial x^3} - E_2\frac{\partial U}{\partial x}\frac{\partial^2 U}{\partial x^2} + E_1\left(\frac{\partial U}{\partial x}\right)^3\right] + \cdots \tag{2.97}$$

$$E_3 = \frac{U}{6}(\alpha^2U^2-1),\quad E_1 = \frac{\alpha}{4}[2(2-\delta)\alpha U + 2r - 1]$$

$$E_2 = -\frac{1}{2}[(r-\delta)\alpha^2U^2 + (2r-1)\alpha U r(1-r)/\delta - 1]$$

不论 r, δ 为何值，上述两步显式守恒格式的截断误差中均有三阶弥散，从而必有三阶相位误差，由此引起的振荡是不可避免的。不同格式 E_2 值不一样，通常要求 $E_2\dfrac{\partial U}{\partial x}<0$ 才保证稳定，且对弥散引起的振荡有压制作用，但对有物理耗散项的控制方程应用此格式要

求 $\left|E_2\frac{\partial U}{\partial x}\right| \ll \bar{\gamma}$，$\bar{\gamma}$ 为物理耗散项前的系数。同时可以看到，对线化问题，$U=a=\text{const}$，$\frac{\partial U}{\partial x}=0$，只存在 $E_3\frac{\partial^3 U}{\partial x^3}$，这与我们对模型方程的分析结果是一致的。一阶导数的三次项对解影响不大，但非线性问题中由于 E_2 的存在，上述二阶格式并不是严格二阶精度的。

通过上述分析可以看出，在非线性条件下，我们按线性分析得到的结论作为参考可以分析问题的性质。

2.9 激波的捕捉法

求解可压流场时可能会出现激波，这是计算流体力学的难题之一。

2.9.1 激波的求解法

一、激波拟合法(Shock Fitting)

基本思想是，在可能出现激波的局部区域内，假定一激波位置，先由解域上游边界开始求到激波前参数的值，于是垂直于激波波线的速度 v_n 与平行于波线的速度 v_s 可解出。由斜激波关系式知，

$$v_{1s}(\text{波前}) = v_{2s}(\text{波后})$$

$$v_{1n}v_{2n} = a_{*n}^2$$

同时用 Rankine-Hugoniot 关系式解出波后的其他物理量，然后由波后的值利用定解条件一直解得全流场值。如果激波位置定得合理，上述解收敛；如不正确则要修正激波位置重复上述过程。

优点：结果准确，不会有跳动和抹平现象；

缺点：位置不易确定，收敛慢，故不常用。

二、激波捕捉法(Shock Capturing)

由守恒的控制方程出发，采用合理的差分格式由定解条件求解，在物理量发生突跃的地方，既算出了突跃量，又确定了激波的位置，收敛较快，计算也较拟合法方便，但有两个明显的缺点：一是激波被抹平(smearing)：真实激波十分薄，至多只有若干个分子平均自由程，只可能在一个网格间达到物理量突变，但数值计算计算时物理量常在 3～5 个网格中变化；二是激波前后解出现振荡。前者主要是数值耗散过度引起的，后者由数值频散引起，也与物理上熵增原理的限制相关。如何从方法上提高激波捕捉的质量是目前高速流体力学计算中较重要的课题之一。本书主要介绍激波捕捉法。

2.9.2 激波捕捉法的一般叙述

1950 年，Von Neumann 与 Richtmyer 开始研究简单方程的解突跃问题：

$$\frac{\partial U}{\partial t} + \frac{\partial (U^2/2)}{\partial x} = 0$$

他们发现，如果用一阶精度的 FTBS(时间前差、空间后差格式)，将解出图 2.12 所示的“S”形的解，抹平程度很大，约 10 个网格，如在多维流中激波将无法辨别。从数值分析观

点来看，主要由于线化修正方程中较大的 $D_2 \frac{\partial^2 U}{\partial x^2}$ 项，这不是物理耗散，而是过大的数值耗散。

如果用二阶精度的格式，如 Lax-Wendroff 格式，则抹平少，但波前后有跳动。这是因为上述流场可看成是 $\frac{\partial U}{\partial t} + \frac{\partial (U^2/2)}{\partial x} = \bar{\gamma}\frac{\partial^2 U}{\partial x^2}$ 中 $\bar{\gamma} \to 0$ 的情况。因此，$Re(\triangle x) = \frac{U}{\bar{\gamma}}\triangle x \to \infty$。修正方程中无二阶耗散，而 4 阶耗散不足以克服三阶弥散，从而在物理量突变处引起解的振荡。为了减少振荡，可以在激波区加上人工的二阶粘性项，此时激波前后振荡减少，抹平约 3～5 个网格。因为二阶精度以上的差分格式，弥散项总是存在，而耗散项不足以完全压制弥散的影响，所以振荡是不同程度存在的。

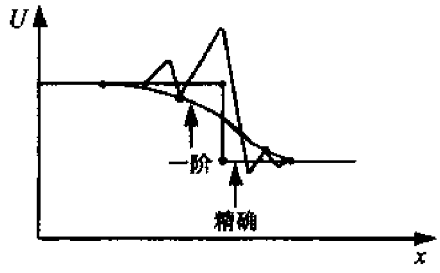


图 2.12

2.10　提高捕捉激波质量的方法

2.10.1　人工粘性滤波法

一阶精度格式，数值粘性过分使激波抹平。如用 MacCormack 等二阶精度的格式，则出现激波前后解的振荡。为消除在激波前后的振荡，20 世纪 70 年代的一些计算中均在激波附近增加人工粘性项以压制振荡，而在物理量无突跃的地区(弥散影响不太大)，此人工粘性项趋于零。

例　用 MacCormack FB 格式解 $\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = 0$。

$$\begin{cases} \overline{U_j^{n+1}} = U_j^n - \frac{\triangle t}{\triangle x}(F_{j+1}^n - F_j^n) + \frac{Q}{2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n) \\ U_j^{n+1} = \frac{1}{2}\left[U_j^n + \overline{U_j^{n+1}} - \frac{\triangle t}{\triangle x}(\overline{F_j^{n+1}} - \overline{F_{j-1}^{n+1}})\right] + \frac{Q}{2}(\overline{U_{j+1}^{n+1}} - 2\overline{U_j^{n+1}} + \overline{U_{j-1}^{n+1}}) \end{cases} \tag{2.98}$$

其中 Q 是开关函数。1978 年 Harten 提出：$Q = \frac{|P_{j+1}^n - P_j^n| - |P_j^n - P_{j-1}^n|}{|P_{j+1}^n - P_j^n| + |P_j^n - P_{j-1}^n|}$；MacCormack 提出：

$$Q = \frac{|P_{j+1}^n - 2P_j^n + P_{j-1}^n|}{|P_{j+1}^n + 2P_j^n + P_{j-1}^n|} \tag{2.99}$$

如激波在 j 与 $j+1$ 之间，则 $P_{j+1}^n \gg P_j^n$；如激波在 j 与 $j-1$ 之间，则 $P_j^n \gg P_{j-1}^n$。由(2.99)知 Q 为有限值。二阶人工粘性项的出现使激波附近格式呈一阶精度。人工粘性项的加入是一项技巧性很强的工作，它使激波不过分抹平，却又能适当地压制弥散引起的振荡。通常无激波光滑区 Q 趋于零，故人工粘性影响较小。

必须指出，在计算每一步所加的人工粘性项，又称光滑项有不同的加法，且对解的稳定性会有一定的影响。

1. 人工粘性光滑项的加法

考虑双曲型方程组：

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}=0$$

(1) 同时光滑法

$$U^{n+1}=(I+R+S)U^{n} \tag{2.100}$$

式中 I 是单位算子，R 是对流项算子，S 是光滑项算子。

以 MacCormack FB 格式预测步为例，并设 $A=\frac{\partial F}{\partial U}$=const，且△为前差算子，则：

$$\begin{aligned}U_j^{\overline{n+1}} &=(I-A\triangle+S)U^{n}\\ &=U_j^n-A\frac{U_{j+1}^n-U_j^n}{\triangle x}\triangle t+\frac{Q}{2}(U_{j+1}^n-2U_j^n+U_{j-1}^n)\\ &=U_j^n-\frac{\triangle t}{\triangle x}(F_{j+1}^n-F_j^n)+\frac{Q}{2}(U_{j+1}^n-2U_j^n+U_{j-1}^n)\end{aligned}$$

这就是我们前面的人工粘性法。它在校正步前就对变量进行光滑，校正步后同样光滑，效果稍差，有时可能会对稳定性起不良影响。

(2) 后置光滑法

$$U^{n+1}=(I+R)U^{n}+S[(I+R)U^{n}] \tag{2.101}$$

可分两步进行：

$$U^{*}=(I+R)U^{n}$$
$$U^{n+1}=(I+S)U^{*}$$

它并非对 U^n 进行光滑运算，而是只对 U^* 进行光滑运算。

如对 MacCormack 格式：

$$U_j^{\overline{n+1}}=U_j^n-\frac{\triangle t}{\triangle x}(F_{j+1}^n-F_j^n) \tag{2.102a}$$

$$U_j^{\overline{\overline{n+1}}}=\frac{1}{2}\left[U_j^n+U_j^{\overline{n+1}}-\frac{\triangle t}{\triangle x}(F_j^{\overline{n+1}}-F_{j-1}^{\overline{n+1}})\right] \tag{2.102b}$$

$$U_j^{n+1}=U_j^{\overline{\overline{n+1}}}+\frac{Q}{2}\left(U_{j+1}^{\overline{\overline{n+1}}}-2U_j^{\overline{\overline{n+1}}}+U_{j-1}^{\overline{\overline{n+1}}}\right) \tag{2.102c}$$

(3) 分裂光滑法

$$U^{*}=(I+R)U^{n}$$
$$U^{**}=(I+S_x)U^{*}$$
$$U^{n+1}=(I+S_y)U^{**}$$

2. 光滑项对稳定性的影响

通过对差分方程用 Von Neumann 分析稳定性的方法计算出放大因子 G 进行分析。以 MacCormack 格式的后置光滑法(2.102)为例，令 $F=AU$，$\sigma=A\frac{\triangle t}{\triangle x}$ 则：

$$G_1=U_j^{\overline{n+1}}/U_j^n=1-\sigma(e^{i\beta h}-1)$$

$$G_2=U_j^{\overline{\overline{n+1}}}/U_j^{\overline{n+1}}=\frac{1}{2}\left[\frac{1}{G_1}+1-\sigma(1-e^{-i\beta h})\right]$$

$$G_3 = U_j^{n+1}/U_j^{\overline{n+1}} = 1 + \frac{Q}{2}(e^{i\beta h} - 2 + e^{-i\beta h})$$

所以

$$G = U_j^{n+1}/U_j^n = G_1G_2G_3 = [1 + Q(\cos\beta h - 1)] \cdot \frac{1}{2}$$

$$\left[1 - \sigma(1 - e^{i\beta h}) + \frac{1}{1 - \sigma(e^{i\beta h} - 1)}\right] \tag{2.103}$$

图 2.13 表示在不同 βh 情况下，放大因子模 $|G|$ 的变化。当 $|G| \to 1.0$ 时，表示耗散对弥散阻尼较弱而趋近于零；βh 在 π 与 $\pi/2$ 之间表示高频频散分量；βh 在 0 与 $\pi/2$ 之间表示低频频散分量。显然，高频分量产生较强的阻尼，使参数光滑，且后光滑的方法效果较佳。图 2.14 表示 σ 与 Q 的稳定区界限。在同时光滑下，稳定区界限比 C.F.L. 条件稍小，在 $Q=0.3$ 时下降更多；而后光滑下在 $Q=0.1\sim0.3$ 时增加了稳定性，使限制放宽。显然，对一维计算而言，后置光滑法可能效果较佳。

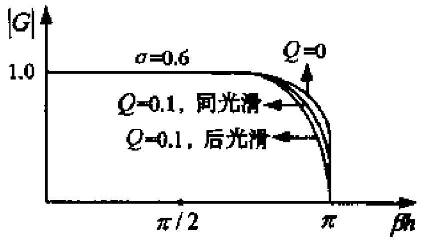


图 2.13

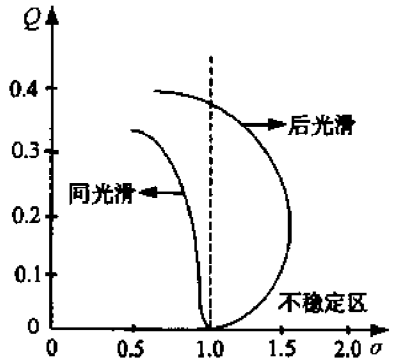


图 2.14

3. 光滑项对守恒性的影响

以质量方程为例，采用同时光滑，一步格式有：

$$\rho_j^{n+1} - \rho_j^n = -\left(\frac{\Delta t}{\Delta x}\right)_j \left[(\rho u)_{j+\frac{1}{2}}^n - (\rho u)_{j-\frac{1}{2}}^n\right] + \frac{Q_j}{2}(\rho_{j+1}^n - 2\rho_j^n + \rho_{j-1}^n)$$

当达到定常 $\rho_j^{n+1} \approx \rho_j^n$ 且收敛时有：

$$(\rho u)_{j+\frac{3}{2}} - (\rho u)_{j+\frac{1}{2}} = -\left(\frac{\Delta x}{2\Delta t}\right)_{j+1} Q_{j+1}(\rho_{j+2} - 2\rho_{j+1} + \rho_j)$$

$$(\rho u)_{j+\frac{1}{2}} - (\rho u)_{j-\frac{1}{2}} = -\left(\frac{\Delta x}{2\Delta t}\right)_j Q_j(\rho_{j+1} - 2\rho_j + \rho_{j-1})$$

$$(\rho u)_{j-\frac{1}{2}} - (\rho u)_{j-\frac{3}{2}} = -\left(\frac{\Delta x}{2\Delta t}\right)_{j-1} Q_{j-1}(\rho_j - 2\rho_{j-1} + \rho_{j-2})$$

……

方程两边求和后，左侧只剩边界值，而方程右侧因为 $Q_{j+1} \neq Q_j \neq Q_{j-1}$，所以解区内有误差积累使守恒性破坏。

如差分格式为：

$$\rho_j^{n+1} - \rho_j^n = -\left(\frac{\Delta t}{\Delta x}\right)_j \cdot \left[(\rho u)_{j+\frac{1}{2}}^n - \frac{Q_{j+\frac{1}{2}}}{2}(\rho_{j+1}^n - \rho_j^n) - (\rho u)_{j-\frac{1}{2}}^n + \frac{Q_{j-\frac{1}{2}}}{2}(\rho_j^n - \rho_{j-1}^n)\right]$$

当达到定常解且收敛时有：

$$(\rho u)_{j+\frac{3}{2}} - (\rho u)_{j+\frac{1}{2}} = \frac{1}{2}Q_{j+\frac{3}{2}}(\rho_{j+2} - \rho_{j+1}) - \frac{1}{2}Q_{j+\frac{1}{2}}(\rho_{j+1} - \rho_j)$$

$$(\rho u)_{j+\frac{1}{2}} - (\rho u)_{j-\frac{1}{2}} = \frac{1}{2}Q_{j+\frac{1}{2}}(\rho_{j+1} - \rho_j) - \frac{1}{2}Q_{j-\frac{1}{2}}(\rho_j - \rho_{j-1})$$

……

当方程两边分别求和时，由于方程右侧对应项能相互抵消而保持了守恒性。

2.10.2 混合格式开关法

上述人工粘性光滑法消除非线性振荡和激波的频散振荡是较为有效的，但在激波附近仍有振荡出现，且在强激波时抹平较严重，有时达5个网格。1976年，Bean-Warming提出提高捕捉激波质量的方法，即开关法。

原理：激波前后的振荡是二阶精度格式的数值弥散在物理量突跃处的反应。模型方程数值试验表明相位误差 $q_F<1$，或 $\phi_F/\phi_e<1$（滞后相位误差），对应的修正方程截断误差项 $D_3<0$（D_3 为修正方程右侧三阶空间导数前的系数）；$q_F>1$，或 $\phi_F/\phi_e>1$（超前相位误差），对应的修正方程截断误差项 $D_3>0$。滞后误差的差分格式在间断后有明显的数值振荡；超前误差在间断前有振荡。我们在间断前用MacCormack格式，它与Lax-Wendroff格式一样，$q_F<1$，波前是光滑的，波后有振荡；而在间断后用Warming-Beam格式，它的 $q_F>1$，在间断后是光滑的，从而保证整个解是光滑的。

考虑

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{u^2}{2}\right) = 0, \quad u(x,0) = \begin{cases} 1, & x < x_0 \\ 0, & x > x_0 \end{cases}$$

1. 采用Lax-Wendroff格式

其修正方程为：

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{u^2}{2}\right) = -\frac{uh^2(1-\alpha^2u^2)}{6}\frac{\partial^3 u}{\partial x^3} + \frac{\alpha^2u^2h^4(\alpha^2u^2-1)}{8k}\frac{\partial^4 u}{\partial x^4} + \cdots \quad (2.104)$$

其中 $\alpha = k/h$。令 $\sigma = u\alpha$，则：

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{u^2}{2}\right) = -\frac{uh^2}{6}(1-\sigma^2)\frac{\partial^3 u}{\partial x^3} + \frac{\alpha^2u^2}{8k}h^4(\sigma^2-1)\frac{\partial^4 u}{\partial x^4} + \cdots$$

显然在稳定性限制条件下，$D_3 = -\frac{uh^2}{6}(1-\sigma^2) < 0, q_F < 1$，所以是滞后相位误差，在波后出现振荡（见图2.15）。

2. Warming-Beam格式

$$u_j^{n+1} = u_j^n - \sigma(u_j^n - u_{j-1}^n) - \frac{\sigma}{2}(1-\sigma)(u_j^n - 2u_{j-1}^n + u_{j-2}^n) \quad (2.105)$$

可证明其修正方程中 $D_3 = \frac{uh^2}{6}(1-\sigma)(2-\sigma)$，当 $\sigma<1$ 时 $D_3>0$，$q_F>1$ 的，所以是波前振荡波后光滑的。具体求解此问题尚必须注意两个问题：

① Warming-Beam 格式是迎风格式，因此在波后使用此格式时，应根据波后流动的特征值的正负号来决定其迎风的方向；

② 激波是弱解，必须保持格式的守恒性。所以在激波前后差分格式改变形式处，使其平均流率满足守恒要求，即加一个开关函数达到此要求。

计算表明，一维流动用此法效果甚佳，但二维流动情况并不太理想。

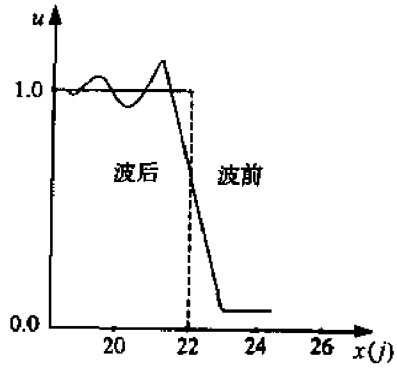


图 2.15

2.10.3 矢通量分裂格式

解双曲型方程组时才可能出现间断性的弱解，而此方程组又要求差分格式必须服从信息传播的方向，即特征方向问题。前面 2.5 节介绍的分离系数矩阵法就反映了这个方向性的问题。但是，用分离系数矩阵的方法不能保证差分方程的守恒，因为一维 Euler 方程组用 $\frac{\partial u}{\partial t} + A\frac{\partial U}{\partial x} = 0$ 表示不是守恒型的方程，捕捉激波有一定困难。

矢通量分裂法就是将守恒型双曲方程组的矢通量 F，按其特征值正负进行分裂，使得在亚声速区内对正特征值对应的 F^+ 用后差格式，对负特征值对应的 F^- 用前差格式，提高精度，并放宽稳定性限制，从而较好地捕捉激波。

对于一维非定常 Euler 方程，可以证明：

$$F(U) = \frac{\partial F}{\partial U}U = A(U)U \tag{2.106}$$

因为 A 有三个互异的实特征值，故

$$\Lambda = P^{-1}AP,\ \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} u & 0 & 0 \\ 0 & u+a & 0 \\ 0 & 0 & u-a \end{bmatrix}。$$

令 $\lambda_i^+ = (\lambda_i + |\lambda_i|)/2, \lambda_i^- = (\lambda_i - |\lambda_i|)/2$，则对任一特征值，有：$\lambda_i = \lambda_i^+ + \lambda_i^-$。

显然，当 $\lambda_i > 0$ 时，$\lambda_i^+ = \lambda_i, \lambda_i^- = 0$；当 $\lambda_i < 0$ 时，$\lambda_i^- = \lambda_i, \lambda_i^+ = 0$。所以，

$$\Lambda = \Lambda^+ + \Lambda^-, \Lambda^+ = \begin{bmatrix} \lambda_1^+ & 0 & 0 \\ 0 & \lambda_2^+ & 0 \\ 0 & 0 & \lambda_3^+ \end{bmatrix}, \Lambda^- = \begin{bmatrix} \lambda_1^- & 0 & 0 \\ 0 & \lambda_2^- & 0 \\ 0 & 0 & \lambda_3^- \end{bmatrix}。$$

亚声速流动时，

$$A^+ = \begin{bmatrix} u & 0 & 0 \\ 0 & u+a & 0 \\ 0 & 0 & 0 \end{bmatrix},\ A^- = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & u-a \end{bmatrix} \tag{2.107}$$

超声速流动时，

$$A^+ = \begin{bmatrix} u & 0 & 0 \\ 0 & u+a & 0 \\ 0 & 0 & u-a \end{bmatrix},\ A^- = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.108}$$

对于方程组(2.80),有:

$$A=\begin{bmatrix}0 & 1 & 0\\ (\gamma-3)\dfrac{u^2}{2} & (3-\gamma)u & \gamma-1\\ (\gamma-1)u^3-\gamma Eu & -3(\gamma-1)\dfrac{u^2}{2}+\gamma E & \gamma u\end{bmatrix} \tag{2.109}$$

右特征矢量为:

$$P=\begin{bmatrix}1 & 1 & 1\\ u & (u+a) & (u-a)\\ \dfrac{u^2}{2} & \left(\dfrac{u^2}{2}+ua+\dfrac{a^2}{\gamma-1}\right) & \left(\dfrac{u^2}{2}-ua+\dfrac{a^2}{\gamma-1}\right)\end{bmatrix} \tag{2.110}$$

$$P^{-1}=\begin{bmatrix}1-\dfrac{u^2}{2}(\gamma-1)a^{-2} & (\gamma-1)ua^{-2} & -(\gamma-1)a^{-2}\\ \beta\left(\dfrac{\gamma-1}{2}u^2-ua\right) & \beta[a-(\gamma-1)u] & \beta(\gamma-1)\\ \beta\left(\dfrac{\gamma-1}{2}u^2+ua\right) & -\beta[a+(\gamma-1)u] & \beta(\gamma-1)\end{bmatrix} \tag{2.111}$$

式中$\beta=1$。

由(2.106),

$$\begin{aligned}F(U)&=AU=P\Lambda P^{-1}U=P(\Lambda^{+}+\Lambda^{-})P^{-1}U=P\Lambda^{+}P^{-1}U+P\Lambda^{-}P^{-1}U\\&=F^{+}(U)+F^{-}(U)\end{aligned}$$

或者

$$F(U)=A^{+}U+A^{-}U=F^{+}(U)+F^{-}(U)$$

微分方程(2.80)可简化为:

$$\frac{\partial U}{\partial t}+\frac{\partial F^{+}(U)}{\partial x}+\frac{\partial F^{-}(U)}{\partial x}=0 \tag{2.112}$$

亚声速下:

$$F^{+}=P\Lambda^{+}P^{-1}U=\frac{\rho}{2\gamma}\begin{bmatrix}(2\gamma-1)u+a\\ 2(\gamma-1)u^2+(u+a)^2\\ (\gamma-1)u^3+\dfrac{(u+a)^3}{2}+\dfrac{(3-\gamma)(u+a)a^2}{2(\gamma-1)}\end{bmatrix} \tag{2.113a}$$

$$F^{-}=P\Lambda^{-}P^{-1}U=\frac{\rho}{2\gamma}\begin{bmatrix}u-a\\ (u-a)^2\\ \dfrac{(u-a)^3}{2}+\dfrac{(3-\gamma)(u-a)a^2}{2(\gamma-1)}\end{bmatrix} \tag{2.113b}$$

超声速下:

$$F^{+}=\begin{bmatrix}\rho u\\ p+\rho u^2\\ (\rho E+p)u\end{bmatrix},\quad F^{-}=0 \tag{2.114}$$

常用的 Warming-Beam 矢通量分裂两步格式为:

$$U_j^{\overline{n+1}} = U_j^n - \frac{\Delta t}{\Delta x}(F_j^{+n} - F_{j-1}^{+n}) - \frac{\Delta t}{\Delta x}(F_{j+1}^{-n} - F_j^{-n}) \tag{2.115a}$$

$$\begin{aligned} U_j^{n+1} = U_j^n &- \frac{\Delta t}{2\Delta x}(F_j^{+n} - F_{j-1}^{+n}) - \frac{\Delta t}{2\Delta x}(F_{j+1}^{-n} - F_j^{-n}) \\ &- \frac{\Delta t}{2\Delta x}(F_j^{+\overline{n+1}} - F_{j-1}^{+\overline{n+1}}) - \frac{\Delta t}{2\Delta x}(F_{j+1}^{-\overline{n+1}} - F_j^{-\overline{n+1}}) \\ &- \frac{\Delta t}{2\Delta x}(F_j^{+n} - 2F_{j-1}^{+n} + F_{j-2}^{+n}) \\ &- \frac{\Delta t}{2\Delta x}(F_{j+2}^{-n} - 2F_{j+1}^{-n} + F_j^{-n}) \end{aligned} \tag{2.115b}$$

其中

$$F_j^{+\overline{n+1}} = F^+(U_j^{\overline{n+1}}) = A^+ U_j^{\overline{n+1}}$$

这样既保证了格式的守恒性，又保证了双曲方程信息传播的方向性，能较好地解出激波。但有些问题需要解决，如F^+，F^-在声速点和驻点附近不光滑的变化会导致解的振荡。这时可用其他方法分解，只须保证$\frac{\partial F^+}{\partial U}$，$\frac{\partial F^-}{\partial U}$分别只有非负和非正的特征值即可。例如，Van Leer 指出，

$$A^+ = \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{bmatrix}, \ A^- = (A - A^+) \tag{2.116}$$

其中

$$\mu_1 = -\frac{-[(u-a)^2 - a^2(\gamma+1)](u+a)^2}{4a^3(\gamma+1)}$$

$$\mu_2 = -\frac{(u-a)[(\gamma-1)u + 2a](u+a)^2}{4a^3(\gamma+1)}$$

$$\mu_3 = \frac{-[(\gamma-1)u^2 + (1-3\gamma)ua + 2(2\gamma+1)a^2](u+a)^2}{4a^3(\gamma+1)}$$

这能保证$A^+ = \frac{\partial F^+}{\partial U}$有非负的特征值，$A^- = \frac{\partial F^-}{\partial U}$有非正的特征值，且这样求得的$F^+$，$F^-$是光滑函数。这说明将矢通量分裂的方法有许多种，分裂后的F^+，F^-也不完全一样。

矢通量分裂格式在求解可压流 Euler 方程组时，捕捉激波的质量是不错的。

2.10.4 总变量递减格式(TVD)与 NND 格式

这是近年来发展的新格式，目的是从数学分析的角度出发提出一种既具二阶精度，又具较小振荡的捕捉激波效果好的差分格式。

1. NND 格式

在介绍 TVD 格式前，先对张涵信提出的一种易于接受，使激波正确捕捉，又较小振荡的所谓 Non-Oscillatory Non-Free-Parameter Dissipation Scheme(简称 NND 格式)进行分析。

考虑双曲型守恒方程 $\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0$，并采用矢通量分裂法：

$$\frac{\partial U}{\partial t}+\frac{\partial F^+}{\partial x}+\frac{\partial F^-}{\partial x}=0 \tag{2.112}$$

可以证明，当 $\frac{\partial F^+}{\partial x}=\frac{1}{2\triangle x}(3F_j^{+n}-4F_{j-1}^{+n}+F_{j-2}^{+n})$（二阶迎风），用半离散格式（$\frac{\partial}{\partial t}$ 项不离散）得，修正方程右端三阶弥散项系数 $D_3>0$，四阶耗散项系数 $D_4<0$。若用中心差分半离散，即 $\frac{\partial F^+}{\partial x}=\frac{1}{2\triangle x}(F_{j+1}^{+n}-F_{j-1}^{+n})$，则 $D_3<0,D_4=0$。结论：在激波附近的物理量，前者是波前振荡，波后光滑；后者正好相反。

同理，如用二阶迎风格式离散 $\frac{\partial F^-}{\partial x}$，则 $D_3<0,D_4<0$。如用二阶中心差分离散 $\frac{\partial F^-}{\partial x}$，则 $D_3>0,D_4=0$。结论：在激波附近的物理量，二阶迎风格式是波前光滑，波后振荡；二阶中心差分正好相反。

上述结论可总结为表 2.2。

表 2.2

$\frac{\partial F^+}{\partial x}$		$\frac{\partial F^-}{\partial x}$	
二阶迎风	二阶中心	二阶迎风	二阶中心
$D_3>0,D_4<0$ 波后光滑	$D_3<0,D_4=0$ 波前光滑	$D_3<0,D_4<0$ 波前光滑	$D_3>0,D_4=0$ 波后光滑

因为我们感兴趣的是在 $t\to\infty$，解趋于定常时的情况，因此采用半离散方法，不计 $\frac{\partial}{\partial t}$ 的影响是可以的，相对也简单些。前面已经分析过 $D_3<0$ 的差分格式，相当于滞后相位误差 $q_F<1$，必在波后即激波的左侧有振荡，而 $D_3>0$ 的格式激波前有振荡。

在对方程(2.112)离散进行激波捕捉时，对 $\frac{\partial F^+}{\partial x}$，我们在激波的左侧（即波后）用二阶迎风差分以保证 $D_3>0$；在激波右侧（即波前）用中心差分保证 $D_3<0$。而对 $\frac{\partial F^-}{\partial x}$ 的波左侧用二阶中心差分保证 $D_3>0$，右侧用二阶迎风差分保证 $D_3<0$。根据表 2.2，这将保证当地物理量光滑。这实质上是混合格式开关法的一种，可表示为：

左侧（波后）：

$$\left(\frac{\partial U}{\partial t}\right)_L=-\left(\frac{\partial F^+}{\partial x}+\frac{\partial F^-}{\partial x}\right)=-\frac{3F_j^{+n}-4F_{j-1}^{+n}+F_{j-2}^{+n}}{2\triangle x}-\frac{F_{j+1}^{-n}-F_{j-1}^{-n}}{2\triangle x} \tag{2.117a}$$

右侧（波前）：

$$\left(\frac{\partial U}{\partial t}\right)_R=-\frac{-3F_j^{-n}+4F_{j+1}^{-n}-F_{j+2}^{-n}}{2\triangle x}-\frac{F_{j+1}^{+n}-F_{j-1}^{+n}}{2\triangle x} \tag{2.117b}$$

将上式写成守恒形式：

$$\frac{\partial U}{\partial t}=-\frac{1}{\triangle x}\left(F_{j+\frac{1}{2}}^n-F_{j-\frac{1}{2}}^n\right) \tag{2.118}$$

$$F_{j+\frac{1}{2}}^n=F_{(j+\frac{1}{2})}^+ + F_{(j+\frac{1}{2})}^-,\quad F_{j-\frac{1}{2}}^n=F_{(j-\frac{1}{2})}^+ + F_{(j-\frac{1}{2})}^-$$

为使(2.118)式的写法能满足(2.117)式，则有：

$$F^{+}_{(j+\frac{1}{2})}=\begin{cases}F_j^{+}+\frac{1}{2}(F_{j+1}^{+}-F_j^{+})=F_j^{+}+\frac{1}{2}\triangle F^{+}_{j+\frac{1}{2}} & (\text{左侧 } L)\\ F_j^{+}+\frac{1}{2}(F_j^{+}-F_{j-1}^{+})=F_j^{+}+\frac{1}{2}\triangle F^{+}_{j-\frac{1}{2}} & (\text{右侧 } R)\end{cases} \tag{2.119}$$

$$F^{-}_{(j+\frac{1}{2})}=\begin{cases}F_{j+1}^{-}-\frac{1}{2}(F_{j+1}^{-}-F_j^{-})=F_{j+1}^{-}-\frac{1}{2}\triangle F^{-}_{j+\frac{1}{2}} & (\text{左侧 } L)\\ F_{j+1}^{-}-\frac{1}{2}(F_{j+2}^{-}-F_{j+1}^{-})=F_{j+1}^{-}-\frac{1}{2}\triangle F^{-}_{j+\frac{3}{2}} & (\text{右侧 } R)\end{cases} \tag{2.120}$$

$$F^{+}_{(j-\frac{1}{2})}=\begin{cases}F_{j-1}^{+}+\frac{1}{2}(F_{j-1}^{+}-F_{j-2}^{+})=F_{j-1}^{+}+\frac{1}{2}\triangle F^{+}_{j-\frac{3}{2}} & (\text{左侧 } L)\\ F_{j-1}^{+}+\frac{1}{2}(F_j^{+}-F_{j-1}^{+})=F_{j-1}^{+}+\frac{1}{2}\triangle F^{+}_{j-\frac{1}{2}} & (\text{右侧 } R)\end{cases} \tag{2.121}$$

$$F^{-}_{(j-\frac{1}{2})}=\begin{cases}F_j^{-}-\frac{1}{2}(F_j^{-}-F_{j-1}^{-})=F_j^{-}-\frac{1}{2}\triangle F^{-}_{j-\frac{1}{2}} & (\text{左侧 } L)\\ F_j^{-}-\frac{1}{2}(F_{j+1}^{-}-F_j^{-})=F_j^{-}-\frac{1}{2}\triangle F^{-}_{j+\frac{1}{2}} & (\text{右侧 } R)\end{cases} \tag{2.122}$$

将(2.119)～(2.122)代入(2.118)得(2.117)，因此是守恒格式。为了在激波左右侧用统一的数学式，我们进一步对 $\triangle F^{+}_{j\pm\frac{1}{2}}$，$\triangle F^{-}_{j\pm\frac{1}{2}}$，$\triangle F^{+}_{j\pm\frac{3}{2}}$，$\triangle F^{-}_{j\pm\frac{3}{2}}$ 的正负及大小进行分析。在二阶精度范围内：

$$\triangle F^{+}_{j\pm\frac{1}{2}}=A_j^{+}\triangle U_{j\pm\frac{1}{2}} \tag{2.123a}$$

$$\triangle F^{-}_{j+1\pm\frac{1}{2}}=A_{j+1}^{-}\triangle U_{j+1\pm\frac{1}{2}} \tag{2.123b}$$

对图 2.16 中单调变化的激波，$\triangle U_{j\pm\frac{1}{2}}$ 在左侧符号相同，在右侧亦如此。在激波的左侧 $\left|\triangle U_{j+\frac{1}{2}}\right|>\left|\triangle U_{j-\frac{1}{2}}\right|$，在激波的右侧 $\left|\triangle U_{j+\frac{1}{2}}\right|<\left|\triangle U_{j-\frac{1}{2}}\right|$。所以，$\triangle F^{+}_{j\pm\frac{1}{2}}$ 的符号满足：

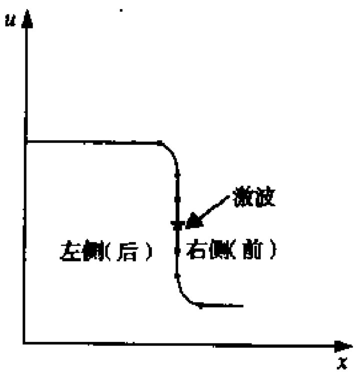


图 2.16

激波左侧：

$$\left|\triangle F^{+}_{j+\frac{1}{2}}\right|>\left|\triangle F^{+}_{j-\frac{1}{2}}\right|,$$

激波右侧：

$$\left|\triangle F^{+}_{j+\frac{1}{2}}\right|<\left|\triangle F^{+}_{j-\frac{1}{2}}\right|$$

$\triangle F^{-}_{j+1\pm\frac{1}{2}}$ 的符号满足：

激波左侧：

$$\left|\triangle F^{-}_{j+\frac{3}{2}}\right|>\left|\triangle F^{-}_{j+\frac{1}{2}}\right|,$$

激波右侧：

$$\left|\triangle F^{-}_{j+\frac{3}{2}}\right|<\left|\triangle F^{-}_{j+\frac{1}{2}}\right|$$

同理可分析 $\triangle F^{+}_{j-1\pm\frac{1}{2}}$ 和 $\triangle F^{-}_{j\pm\frac{1}{2}}$ 的符号。

取 min mod(x,y)，有下列含义：当 x,y 同号时，取 x,y 中绝对值最小的那个元素；x，

y 异号时取 0。于是有：

$$F^{+}_{(j+\frac{1}{2})} = F^{+}_{j} + \frac{1}{2}\min \bmod(\triangle F^{+}_{j-\frac{1}{2}}, \triangle F^{+}_{j+\frac{1}{2}}) \tag{2.124}$$

$$F^{-}_{(j+\frac{1}{2})} = F^{-}_{j+1} - \frac{1}{2}\min \bmod(\triangle F^{-}_{j+\frac{1}{2}}, \triangle F^{-}_{j+\frac{3}{2}}) \tag{2.125}$$

$$F^{+}_{(j-\frac{1}{2})} = F^{+}_{j-1} + \frac{1}{2}\min \bmod(\triangle F^{+}_{j-\frac{3}{2}}, \triangle F^{+}_{j-\frac{1}{2}}) \tag{2.126}$$

$$F^{-}_{(j-\frac{1}{2})} = F^{-}_{j} - \frac{1}{2}\min \bmod(\triangle F^{-}_{j-\frac{1}{2}}, \triangle F^{-}_{j+\frac{1}{2}}) \tag{2.127}$$

显然，(2.124)～(2.127)分别代表了(2.119)～(2.122)，所以，(2.118)与(2.117)可写成：

$$\begin{aligned} U^{n+1}_{j} =& U^{n}_{j} - \frac{\triangle t}{\triangle x}\cdot \\ & \left\{\left[F^{+}_{j} + \frac{1}{2}\min \bmod(\triangle F^{+}_{j-\frac{1}{2}}, \triangle F^{+}_{j+\frac{1}{2}}) + F^{-}_{j+1} - \frac{1}{2}\min \bmod(\triangle F^{-}_{j+\frac{1}{2}}, \triangle F^{-}_{j+\frac{3}{2}})\right]\right. \\ & \left. - \left[F^{+}_{j-1} + \frac{1}{2}\min \bmod(\triangle F^{+}_{j-\frac{3}{2}}, \triangle F^{+}_{j-\frac{1}{2}}) + F^{-}_{j} - \frac{1}{2}\min \bmod(\triangle F^{-}_{j-\frac{1}{2}}, \triangle F^{-}_{j+\frac{1}{2}})\right]\right\} \end{aligned} \tag{2.128}$$

此格式除个别点为一阶精度外均具有无耗散无跳动的二阶精度。这种格式物理概念较明显，易于接受。以后可以证明这就是 TVD 格式的一种。

2. TVD(Total Variable Diminishing)格式

(1) 概述

用前面各种方法解一维 Euler 方程出现不连续的突跃解时，有三个难点：

① 解一旦不光滑，二阶格式就会出现振荡；

② 常系数情况下稳定的状态在突跃处有非线性不稳定的增长；

③ 差分方程的解可能是非物理解。

通常单调的守恒型格式的收敛解导致满足熵条件的物理弱解，这样上述三个问题都能较好解决，但格式精度是一阶的。下面从一维非线性双曲型方程出发讨论 TVD 格式。

考虑

$$\frac{\partial u}{\partial t} + \frac{\partial F}{\partial x} = 0,\ \frac{\partial F}{\partial u} = A(u) \tag{2.80}$$

在初值 $u(x,0)=g_0(x)$ 情况下解其初值问题(为书写方便，将 U 写成了 u)。这个方程的解可以用 $x-t$ 平面上一系列散开的或收拢的特征线描述。Lax 指出，任何一双特征线间解的总变量的增减是守恒的。即：

$$I(t+\triangle t) = I(t) = I(t_0) \tag{2.129a}$$

$$I(t) = \int_{-\infty}^{+\infty}\left|\frac{\partial u(x,t)}{\partial x}\right|\mathrm{d}x \tag{2.129b}$$

Lax 还指出，在条件(2.129)下，单调的解保持单调。但如果解区中有激波存在，则解的总变量总是下降的，即：

$$I(t+\triangle t) \leqslant I(t) \tag{2.130}$$

这是因为当同族特征线相交形成激波时，解的变动（信息）就会有所损失（不可逆），所以总变量递减。目前只能作如此解释。

当 $u_j^n = u(j\triangle x, n\triangle t)$ 时，若令 $TV(u) = \sum_{-\infty}^{\infty} |u_{j+1} - u_j|$，则(2.130)为：

$$TV(u^{n+1}) \leqslant TV(u^n) \tag{2.131}$$

这被称为 TVD 条件。

Harten(1983)指出，上述 TVD 条件加上格式的差分方程与微分方程相容和满足熵不等式，就能在单调的前提下保证解收敛到弱解。

(2) 构造

下面构造 TVD 显格式。

$$u_j^{n+1} = u_j^n - \sum_{i=-p}^{q} C(i)_{j+\frac{1}{2}} \triangle_{j+\frac{1}{2}} u^n, \triangle_{j+\frac{1}{2}} u = u_{j+1} - u_j \tag{2.132}$$

因为 C 是 u 的非线性函数，因此上式的具体形式不是唯一的，便于我们有选择格式的自由度。现在只介绍构造 TVD 格式的一般原则。

取一个向前差分矩阵 D，$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & -1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & -1 \end{bmatrix}$，并定义 $\| Du \|_1$ 为 Du 所有元素的绝对值之和，则：

$$\| Du \|_1 \equiv TV(u) \tag{2.133}$$

所以，TVD 条件为：

$$\| Du^{n+1} \|_1 \leqslant \| Du^n \|_1 \tag{2.134}$$

格式(2.132)可写成：

$$u^{n+1} = [I - MD]u^n = Ru^n \tag{2.135}$$

M 是矩阵算子，它可以是 u 的非线性函数。由(2.135)得：

$$Du^{n+1} = [I - MD]Du^n = RDu^n$$

所以

$$\| Du^{n+1} \|_1 \leqslant \| R \|_1 \| Du^n \|_1 \tag{2.136}$$

比较上式与(2.134)知，格式(2.132)或(2.135)是 TVD 的充要条件是：

$$\| R \|_1 \leqslant 1 \tag{2.137}$$

因为 R 的 L_1 模是其 R 中每列的元素绝对值的和，很显然，为保证充要条件 $\| R \|_1 \leqslant 1$，必要求 $R>0$，即 R 为非负矩阵。

按照矢通量分裂原则，

$$\begin{aligned} u_j^{n+1} &= u_j^n + C_{j+\frac{1}{2}} \triangle_{j+\frac{1}{2}} u^n - C_{j-\frac{1}{2}} \triangle_{j-\frac{1}{2}} u^n \\ &= u_j^n + C_{j+\frac{1}{2}} (u_{j+1}^n - u_j^n) - C_{j-\frac{1}{2}} (u_j^n - u_{j-1}^n) \\ &= C_{j-\frac{1}{2}} u_{j-1}^n + (1 - C_{j+\frac{1}{2}} - C_{j-\frac{1}{2}}) u_j^n + C_{j+\frac{1}{2}} u_{j+1}^n \end{aligned}$$

此式与(2.135)相比得：

$$R=\begin{bmatrix} 0 & \cdots & 0 & C_{j+\frac{1}{2}} & 0 & \cdots & 0 \\ \cdots & 0 & C_{j-\frac{1}{2}} & 1-C_{j+\frac{1}{2}}-C_{j-\frac{1}{2}} & C_{j+\frac{1}{2}} & 0 & \cdots \\ 0 & \cdots & 0 & C_{j-\frac{1}{2}} & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 0 & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & 0 & \ddots & \ddots & \ddots \end{bmatrix}$$

所以，上述差分格式满足 TVD 要求的充要条件为：

$$C_{j-\frac{1}{2}}\geqslant 0,\ (1-C_{j+\frac{1}{2}}-C_{j-\frac{1}{2}})\geqslant 0,\ C_{j+\frac{1}{2}}\geqslant 0 \tag{2.138}$$

(3) 说明

取(2.138)中 C 的形成的途径是很多的。例如，NND 格式就是 TVD。由(2.129)，令 $0\leqslant k_i\leqslant 1$，$i=1,2,3,4$，则：

$$U_j^{n+1}-U_j^n=-\alpha\left(\Delta F_{j-\frac{1}{2}}^{+n}+\Delta F_{j+\frac{1}{2}}^{-n}\right)-\frac{\alpha k_1}{2}\Delta F_{j-\frac{1}{2}}^{+n}+\frac{\alpha k_3}{2}\Delta F_{j+\frac{1}{2}}^{-n}$$

$$+\frac{\alpha k_2}{2}\Delta F_{j-\frac{1}{2}}^{+n}-\frac{\alpha k_4}{2}\Delta F_{j+\frac{1}{2}}^{-n}$$

$$=-\alpha\left(r\Delta F_{j-\frac{1}{2}}^{+n}+\beta\Delta F_{j+\frac{1}{2}}^{-n}\right)$$

$$r=\left[1-\left(\frac{k_2}{2}-\frac{k_1}{2}\right)\right],\ \beta=1-\frac{k_3-k_4}{2}$$

所以，

$$U_j^{n+1}-U_j^n=-\bar{r}_{j-\frac{1}{2}}\Delta U_{j-\frac{1}{2}}^n+\bar{\beta}_{j+\frac{1}{2}}\Delta U_{j+\frac{1}{2}}^n)\sim C_{j+\frac{1}{2}}^{+}\Delta_{j+\frac{1}{2}}U^n-C_{j-\frac{1}{2}}^{-}\Delta_{j-\frac{1}{2}}U^n$$

即

$$\bar{\beta}_{j+\frac{1}{2}}=C_{j+\frac{1}{2}},\ \bar{r}_{j-\frac{1}{2}}=C_{j-\frac{1}{2}},$$

式中

$$\bar{r}_{j-\frac{1}{2}}=\alpha r\frac{F_j^{+n}-F_{j-1}^{+n}}{U_j^n-U_{j-1}^n}=\alpha r a^{+}=\alpha r\frac{\partial F^{+}}{\partial U}$$

$$\bar{\beta}_{j+\frac{1}{2}}=-\alpha\beta\frac{F_{j+1}^{-n}-F_j^{-n}}{U_{j+1}^n-U_j^n}=-\alpha\beta a^{-}=-\alpha\beta\frac{\partial F^{-}}{\partial U}$$

$$a^{+}\geqslant 0,\ a^{-}\leqslant 0,$$

所以

$$\bar{r}_{j-\frac{1}{2}}\geqslant 0,\ \bar{\beta}_{j+\frac{1}{2}}\geqslant 0$$

即

$$C_{j-\frac{1}{2}}\geqslant 0,\ C_{j+\frac{1}{2}}\geqslant 0$$

同理，在一般情况下，也可证明 $C_{j-\frac{1}{2}}+C_{j+\frac{1}{2}}\leqslant 1$。因此，NND 格式是 TVD 格式。用 NND 解 $\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=\mu\frac{\partial^2 u}{\partial x^2}$，当 $\mu=0.0001$ 时分辨率很高，如图 2.17 所示。

Lax-Wendroff 格式不是 TVD 格式。以标量方程 $\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0$ 为例,用此格式离散

$$
\begin{aligned}
u_j^{n+1}-u_j^n &= -\frac{a\alpha}{2}(u_{j+1}^n-u_{j-1}^n)+\frac{a^2\alpha^2}{2}(u_{j+1}^n-2u_j^n+u_{j-1}^n) \\
&= -\frac{a\alpha}{2}(\triangle u_{j+\frac{1}{2}}^n+\triangle u_{j-\frac{1}{2}}^n)+\frac{a^2\alpha^2}{2}(\triangle u_{j+\frac{1}{2}}^n-\triangle u_{j-\frac{1}{2}}^n) \\
&= C_{j+\frac{1}{2}}\triangle u_{j+\frac{1}{2}}^n-C_{j-\frac{1}{2}}\triangle u_{j-\frac{1}{2}}^n
\end{aligned}
$$

其中

$$C_{j+\frac{1}{2}}=\frac{\alpha}{2}(-a+\alpha a^2)=\frac{\alpha a}{2}(\alpha a-1)$$

$$C_{j-\frac{1}{2}}=\frac{\alpha}{2}(a+\alpha a^2)=\frac{\alpha a}{2}(\alpha a+1)$$

已知稳定性条件为:$a\alpha\leqslant 1$,所以 $C_{j-\frac{1}{2}}\geqslant 0$, $C_{j+\frac{1}{2}}\leqslant 0$,$C_{j-\frac{1}{2}}+C_{j+\frac{1}{2}}\leqslant 1$ 不满足(2.138)式,这说明用 Lax-Wendroff 格式解激波有振荡。

总之,根据 TVD 要求可构成一系列的 TVD 格式。

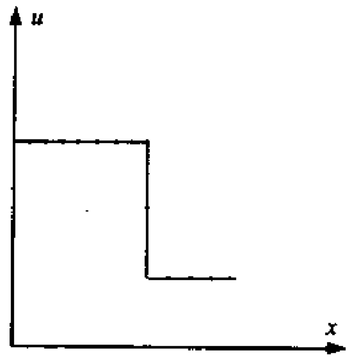


图 2.17

2.10.5 自适应网格法

自适应网格法(Adaptive Grid Methods)不是从差分格式的构成出发,而是从网格生成的途径来提高捕捉激波的分辨率,达到提高精度的要求。它从物理量的梯度出发构成网格。物理量光滑的地区,网格分布稀疏一些,而物理量梯度大的地方,尤其是激波、燃烧波的附近及粘性湍流边界层的附近网格加密。这种网格称自适应网格,详细分析于后文再述。

第三章　二维、三维流动抛物型与双曲型方程差分研究

3.1　抛物型方程显式差分格式

对于微分方程：

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2} \tag{3.1}$$

常见的差分格式为：

$$u_{j,m}^{n+1}=[1+r(\delta_{xx}+\delta_{yy})]u_{j,m}^{n} \tag{3.2}$$

其中 $r=\Delta t/\Delta x^2=\Delta t/\Delta y^2$，$\delta_{xx}$和 δ_{yy}为二阶中心差分。Taylor 级数展开后(3.2)式的截断误差首项的量级是 $O(k,h^2)$，即(3.2)是时间一阶、空间二阶的差分格式。

(3.2)的稳定性极限为$r\leqslant 1/4$。当 $\Delta x\neq\Delta y$ 时，则(3.2)的稳定性极限为 $r_x+r_y\leqslant 1/2$。

推论：对三维流，当差分格式为 $u_{j,m,l}^{n+1}=[1+r(\delta_{xx}+\delta_{yy}+\delta_{zz})]u_{j,m,l}^{n}$时，稳定性极限为：

$$r\leqslant 1/6 \quad (\Delta x=\Delta y=\Delta z)$$

$$r_x+r_y+r_z\leqslant 1/2 \quad (\Delta x\neq\Delta y\neq\Delta z)$$

3.2　双曲型与输运型方程显式差分格式

3.2.1　平流方程

考虑方程

$$\frac{\partial u}{\partial t}+a_1\frac{\partial u}{\partial x}+a_2\frac{\partial u}{\partial y}=0,\quad a_1>0,a_2>0 \tag{3.3}$$

一、一阶迎风格式

$$u_{j,m}^{n+1}=u_{j,m}^{n}-\sigma_1(u_{j,m}^{n}-u_{j-1,m}^{n})-\sigma_2(u_{j,m}^{n}-u_{j,m-1}^{n}) \tag{3.4}$$

其中 $\sigma_1=a_1\cdot\frac{k}{h_x}$，$\sigma_2=a_2\cdot\frac{k}{h_y}$，$h_x$，$h_y$ 分别是 x 和 y 向的网格间距。(3.4)式截断误差首项的量级是 $O(k,h_x,h_y)$，故它是时间、空间均为一阶精度的格式。

(3.4)的稳定性条件为：$\sigma_1+\sigma_2\leqslant 1$。当 $\sigma_1=\sigma_2$ 时，$\sigma_1=\sigma_2\leqslant 1/2$。

对三维情况，稳定性条件为：$\sigma_1+\sigma_2+\sigma_3\leqslant 1$。当 $\sigma_1=\sigma_2=\sigma_3$ 时，$\sigma_1=\sigma_2=\sigma_3\leqslant 1/3$。

二、Lax-Wendroff 格式

$$u_{j,m}^{n+1}=u_{j,m}^{n}+\left(\frac{\partial u}{\partial t}\right)_{j,m}^{n}k+\frac{k^2}{2}\left(\frac{\partial^2 u}{\partial t^2}\right)_{j,m}^{n}+O(k^3) \tag{3.5}$$

(3.5)式中时间导数项均用空间导数项取代，并且对空间导数项仍用中心差分，则：

$$u_{j,m}^{n+1}=\left[1-\frac{\sigma_1}{2}\overline{\delta_x}-\frac{\sigma_2}{2}\overline{\delta_y}+\frac{1}{2}(\sigma_1^2\delta_{xx}+\sigma_2^2\delta_{yy})+\frac{1}{4}\sigma_1\sigma_2\delta_{xy}\right]u_{j,m}^n \tag{3.6}$$

其中 $\overline{\delta_x}$：一阶中心差$\overline{\delta_x}u_{j,m}^n=(u_{j+1,m}^n-u_{j-1,m}^n)$；

δ_{xx},δ_{yy}：二阶中心差；

$\delta_{xy}=\overline{\delta_x}(\overline{\delta_y})$

可以证明，当 $\sigma_1=\sigma_2=\sigma$ 时，$\sigma\leqslant 1/2$。

3.2.2 双曲型方程组的显式差分格式

对于微分方程组：

$$\frac{\partial U}{\partial t}+A\frac{\partial U}{\partial x}+B\frac{\partial U}{\partial y}=0 \tag{3.7}$$

其中 A、B 均为系数矩阵。令 $A_1=-A,B_1=-B$，按(3.6)有以下 Lax-Wendroff 格式：

$$U_{j,m}^{n+1}=\left[I+\frac{\sigma_1}{2}\overline{\delta_x}+\frac{\sigma_2}{2}\overline{\delta_y}+\frac{1}{2}(\sigma_1^2\delta_{xx}+\sigma_2^2\delta_{yy})+\frac{1}{8}\sigma_1'\sigma_2'(A_1B_1+B_1A_1)\delta_{xy}\right]U_{j,m}^n \tag{3.8}$$

其中 $\sigma_1=A_1\cdot\frac{k}{h_x},\sigma_2=B_1\cdot\frac{k}{h_y},\sigma_1'=\frac{k}{h_x},\sigma_2'=\frac{k}{h_y}$，注意：$A_1B_1\neq B_1A_1$.

设 A_1,B_1 的特征值分别为 λ_A,λ_B，则写出差分方程的放大矩阵 G 后，由其特征值 $\mu(G)$ 求出 $\mu_{\max}$，满足 $|\mu|_{\max}\leqslant 1$ 得稳定性极限为：

$$\frac{k}{h}\max\{|\lambda_A|,|\lambda_B|\}\leqslant\frac{1}{2} \tag{3.9}$$

3.2.3 输运型方程的差分格式

考虑方程：

$$\frac{\partial u}{\partial t}+a_1\frac{\partial u}{\partial x}+a_2\frac{\partial u}{\partial y}=\overline{\gamma}_1\frac{\partial^2 u}{\partial x^2}+\overline{\gamma}_2\frac{\partial^2 u}{\partial y^2} \tag{3.10}$$

这是抛物型方程，但与纯抛物型有许多差别，当流场的雷诺数 Re 较大时，(3.10)右侧按其物理意义 γ_1 和 γ_2 很小，方程具有双曲型性质，通常用中心差分：

$$\begin{aligned}u_{j,m}^{n+1}=&(r_x+\frac{\sigma_1}{2})u_{j-1,m}^n+(r_y+\frac{\sigma_2}{2})u_{j,m-1}^n+(1-2r_x-2r_y)u_{j,m}^n\\&+(r_x-\frac{\sigma_1}{2})u_{j+1,m}^n+(r_y-\frac{\sigma_2}{2})u_{j,m+1}^n\end{aligned} \tag{3.11}$$

其中 $r_x=\overline{\gamma}_1\cdot\frac{k}{h_x^2},r_y=\overline{\gamma}_2\cdot\frac{k}{h_y^2},\sigma_1=a_1\cdot\frac{k}{h_x},\sigma_2=a_2\cdot\frac{k}{h_y}$.

稳定性分析得稳定性条件为：$r_x+r_y\leqslant 1/2,\sigma_1+\sigma_2\leqslant 1$。当 $a_1=a_2,h_x=h_y$ 时，$\sigma_1=\sigma_2\leqslant 1/2$。

事实上，上述稳定性条件是二维平流方程与抛物型方程的迭加。这种结果很有启发性，说明在解问题时，对流项与耗散项可分别求解后加以迭加，每一过程均用自己的稳定性限制。

输运型方程组就是类似于 N.S. 方程的方程组。如：

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=\frac{\partial}{\partial x}(\bar{\gamma}_1\frac{\partial U}{\partial x})+\frac{\partial}{\partial y}(\bar{\gamma}_2\frac{\partial U}{\partial y}) \tag{3.12a}$$

即：

$$\frac{\partial U}{\partial t}+\frac{\partial}{\partial x}(F-\bar{\gamma}_1\frac{\partial U}{\partial x})+\frac{\partial}{\partial y}(G-\bar{\gamma}_2\frac{\partial U}{\partial y})=0 \tag{3.12b}$$

采用 MacCormack 格式，耗散项用中心差分，得：

$$U_{j,m}^{\overline{n+1}}=U_{j,m}^n-\frac{k}{h_x}[\triangle_x(F_{j,m}^n-\bar{\gamma}_{1j,m}^n\nabla_x U_{j,m}^n)]$$

$$-\frac{k}{h_y}[\triangle_y(G_{j,m}^n-\bar{\gamma}_{2j,m}^n\nabla_y U_{j,m}^n)] \tag{3.13a}$$

$$U_{j,m}^{n+1}=\frac{1}{2}\Big\{U_{j,m}^n+U_{j,m}^{\overline{n+1}}-\frac{k}{h_x}[\nabla_x(F_{j,m}^{\overline{n+1}}-\bar{\gamma}_{1j,m}^{\overline{n+1}}\triangle_x U_{j,m}^{\overline{n+1}})]$$

$$-\frac{k}{h_y}[\nabla_y(G_{j,m}^{\overline{n+1}}-\bar{\gamma}_{2j,m}^{\overline{n+1}}\triangle_y U_{j,m}^{\overline{n+1}})]\Big\} \tag{3.13b}$$

式中$\triangle_x,\triangle_y$为空间前差，∇_x,∇_y为空间后差。设$\Delta x=\Delta y=h,\bar{\gamma}_1=\bar{\gamma}_2,\lambda_A=\lambda_B=\lambda$，则稳定性估算法为：

$$\frac{1}{\Delta t}=\frac{1}{\Delta t_{无粘}}+\frac{1}{\Delta t_{有粘}}\geqslant\frac{2|\lambda|}{h}+\frac{4\bar{\gamma}}{h^2}$$

所以

$$\Delta t\leqslant\frac{h^2}{4\bar{\gamma}+2|\lambda|h} \tag{3.14}$$

(3.14)中的$\bar{\gamma}$是运动粘性系数。常温下，空气的$\bar{\gamma}$约为1.5×10^{-5}米^2/秒；λ为无粘 Euler 方程的特征值，通常是 300 米/秒的量级，h为$10^{-2}\sim10^{-3}$米量级，因此(3.14)中的Δt主要由$h/2|\lambda|$来确定。

3.2.4 显式时间分裂法——“和”分裂法

用上述显式方法解二维问题稳定性限制较一维流更加严格。尤其在解航空航天工程的实际流动时，由于 Re 较高，湍流边界层是主要的粘性影响，在壁面沿法向十分薄的区间内，物理量变化甚快，因此粘性流区内通常要求$\Delta y\ll\Delta x$。稳定性极限为：

$$\frac{1}{\Delta t_x}=\frac{1}{\Delta t_{x无粘}}+\frac{1}{\Delta t_{x有粘}}\geqslant\frac{|\lambda_{Am}|}{\Delta x}+\frac{2\bar{\gamma}_1}{\Delta x^2}$$

$$\frac{1}{\Delta t_y}=\frac{1}{\Delta t_{y无粘}}+\frac{1}{\Delta t_{y有粘}}\geqslant\frac{|\lambda_{Bm}|}{\Delta y}+\frac{2\bar{\gamma}_2}{\Delta y^2}$$

$$\Delta t_x\leqslant\frac{1}{\frac{|\lambda_{Am}|}{\Delta x}+\frac{2\bar{\gamma}_1}{\Delta x^2}},\Delta t_y\leqslant\frac{1}{\frac{|\lambda_{Bm}|}{\Delta y}+\frac{2\bar{\gamma}_2}{\Delta y^2}}$$

$$\Delta t\leqslant\min(\Delta t_x,\Delta t_y)$$

因为$\Delta t_y\ll\Delta t_x$，所以稳定性限制主要通过Δt_y来反映，时间步长主要因Δy小而大为减小。如在计算$\frac{\Delta t_x}{\Delta x}(F_{j,m}-F_{j-1,m})$时也用同一时间步长$\Delta t_y$，则大大浪费机时。下面设法将二维问题分成两个一维问题分别求解来节省机时，提高效率。

为了说明问题，先从无粘的方程作一个方法上的分析与介绍，再用到粘性流动中去。

考虑方程组：

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \tag{3.15}$$

此方程由方程(1.1)略去粘性项，并将原方程的 F_I, G_I 分别写成 F, G 而得到。

(3.15)的 Lax-Wendroff 格式为：

$$\begin{aligned}U^{n+1}=\Big\{&I+\Delta t\left(A_1\frac{\partial}{\partial x}+B_1\frac{\partial}{\partial y}\right)\\&+\frac{\Delta t^2}{2}\left[A_1^2\frac{\partial^2}{\partial x^2}+B_1^2\frac{\partial^2}{\partial y^2}+(A_1B_1+B_1A_1)\frac{\partial^2}{\partial x\partial y}\right]\Big\}U^n+O(k^3)\end{aligned} \tag{3.16}$$

其中 $A_1=-\frac{\partial F}{\partial U}, B_1=-\frac{\partial G}{\partial U}$。

令一维算子：

$$L_x(\Delta t)U^n=\left[I+\Delta tA_1\frac{\partial}{\partial x}+\frac{\Delta t^2}{2}A_1^2\frac{\partial^2}{\partial x^2}\right]U^n+O(k^3) \tag{3.17a}$$

$$L_y(\Delta t)U^n=\left[I+\Delta tB_1\frac{\partial}{\partial y}+\frac{\Delta t^2}{2}B_1^2\frac{\partial^2}{\partial y^2}\right]U^n+O(k^3) \tag{3.17b}$$

若 $\Delta x=\Delta y$，则：

$$\begin{aligned}[L_y(\Delta t)L_x(\Delta t)]U^n&=\left[I+\Delta tB_1\frac{\partial}{\partial y}+\frac{\Delta t^2}{2}B_1^2\frac{\partial^2}{\partial y^2}\cdot I+\Delta tA_1\frac{\partial}{\partial x}+\frac{\Delta t^2}{2}A_1^2\frac{\partial^2}{\partial x^2}\right]U^n+O(k^3)\\&=\Big\{I+\Delta t\left(A_1\frac{\partial}{\partial x}+B_1\frac{\partial}{\partial y}\right)\\&\quad+\frac{\Delta t^2}{2}\cdot\left[A_1^2\frac{\partial^2}{\partial x^2}+B_1^2\frac{\partial^2}{\partial y^2}+(A_1B_1+B_1A_1)\frac{\partial^2}{\partial x\partial y}\right]\Big\}U^n+O(k^3)\end{aligned}$$

如果 $A_1B_1=B_1A_1$，则：

$$U^{n+1}=[L_y(\Delta t)L_x(\Delta t)]U^n+O(k^3) \tag{3.18}$$

保持二阶精度不变。

令 $\overline{U^{n+1}}=L_x(\Delta t)U^n$，则：

$$U^{n+1}=L_y(\Delta t)\overline{U^{n+1}}$$

每一步均满足 C. F. L. 条件，即：$|\lambda_A|\frac{\Delta t_x}{\Delta x}\leqslant 1$，$|\lambda_B|\frac{\Delta t_y}{\Delta y}\leqslant 1$。又 $\Delta x=\Delta y$，故 $\Delta t\leqslant\frac{\Delta x}{|\lambda_m|}$。

如令 $L_{x1}=A_1\frac{\partial}{\partial x}, L_{y1}=B_1\frac{\partial}{\partial y}$，则前面问题可小结如下：

$$\begin{aligned}U^{n+1}&=\left(\exp k\frac{\partial}{\partial t}\right)U^n\\&=[\exp k(L_{x1}+L_{y1})]U^n\\&=[\exp kL_{x1}\exp kL_{y1})]U^n\\&=\left[\left(I+kL_{x1}+\frac{k^2}{2}L_{x1}^2\right)\left(I+kL_{y1}+\frac{k^2}{2}L_{y1}^2\right)\right]U^n+O(k^3)\\&=L_x(\Delta t)L_y(\Delta t)U^n\end{aligned}$$

这就是(3.18)式，但当 $A_1B_1\neq B_1A_1$ 或 $\Delta y\ll\Delta x$ 时，应采取下列措施：

一、$A_1B_1 \neq B_1A_1$

为保证精度，令

$$
\begin{aligned}
U^{n+1} &= \exp k(L_{x1}+L_{y1})U^n \\
&= \left(\exp\frac{k}{2}L_{y1}\exp kL_{x1}\exp\frac{k}{2}L_{y1}\right)U^n \\
&= \left[L_y(\frac{k}{2})L_x(k)L_y(\frac{k}{2})\right]U^n
\end{aligned}
\tag{3.19}
$$

$$
\begin{aligned}
U^{n+2} &= \exp 2k(L_{x1}+L_{y1})U^n \\
&= (\exp kL_{y1}\exp kL_{x1}\exp kL_{x1}\exp kL_{y1})U^n \\
&= [L_y(k)L_x(k)L_x(k)L_y(k)]U^n
\end{aligned}
$$

二、$\Delta y \ll \Delta x$

因 Re 很大，γ 很小，粘性对稳定性的影响可忽略不计。由 $|\lambda_m|\frac{\Delta t_x}{\Delta x}\leqslant 1$，$|\lambda_m|\frac{\Delta t_y}{\Delta y}\leqslant 1$ 得：$\Delta t_x \leqslant \frac{\Delta x}{|\lambda_m|}$，$\Delta t_y \leqslant \frac{\Delta y}{|\lambda_m|}$。一般取 $M=\frac{\Delta t_x}{2\Delta t_y}$，$M$ 为正整数。如 $\Delta x = 100\Delta y$，则 $\Delta t_x = 100\Delta t_y$，$M=50$，于是：

$$
U^{n+1} = \left[L_y(\frac{\Delta t_x}{2M})^M L_x(\Delta t_x)L_y(\frac{\Delta t_x}{2M})^M\right]U^n
$$

显然，L_x 步数减少，效率提高。稳定性限制放宽也使效率提高。

三、以 $\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}-\frac{\partial}{\partial x}(\bar{\gamma}_1\frac{\partial U}{\partial x})-\frac{\partial}{\partial y}(\bar{\gamma}_2\frac{\partial U}{\partial y})=0$ 的 MacCormack 格式为例

$$
L_y: \overline{U_{j,m}^{n+\frac{1}{2M}}} = U_{j,m}^n - \frac{\Delta t}{2M\Delta y}\Delta_y(G_{j,m}^n - \bar{\gamma}_{1j,m}^n\nabla_y U_{j,m}^n)
$$

$$
U_{j,m}^{n+\frac{1}{2M}} = \frac{1}{2}\left\{U_{j,m}^n + \overline{U_{j,m}^{n+\frac{1}{2M}}} - \frac{\Delta t}{2M\Delta y}\nabla_y(\overline{G_{j,m}^{n+\frac{1}{2M}}} - \overline{\bar{\gamma}_{1j,m}^{n+\frac{1}{2M}}}\Delta_y\overline{U_{j,m}^{n+\frac{1}{2M}}})\right\}
\tag{3.20a}
$$

重复 M 次，得到 $U_{j,m}^{n+\frac{1}{2}}$；

$$
L_x: U_{j,m}^{\overline{n+1}} = U_{j,m}^{n+\frac{1}{2}} + \frac{\Delta t}{\Delta x}\Delta_x(F_{j,m}^{n+\frac{1}{2}} - \bar{\gamma}_{2j,m}^{n+\frac{1}{2}}\nabla_x U_{j,m}^{n+\frac{1}{2}})
$$

$$
U_{j,m}^{\overline{n+1}} = \frac{1}{2}\left\{U_{j,m}^{\overline{n+1}} + U_{j,m}^{n+\frac{1}{2}} - \frac{\Delta t}{\Delta x}\nabla_x(F_{j,m}^{\overline{n+1}} - \bar{\gamma}_{1j,m}^{\overline{n+1}}\Delta_x U_{j,m}^{\overline{n+1}})\right\}
\tag{3.20b}
$$

L_y 再以 $U_{j,m}^{\overline{n+1}}$ 为初值，每次步进$(1/2)M$，总共步进 M 次，与(3.20a)进行类似计算得 $U_{j,m}^{n+1}$。

还有一种解决办法是在 L_y 用隐式，L_x 仍用显式，使 L_y 的 $2M$ 次迭代次数减少，因隐式允许用较大的时间步长，但隐式每步计算较显式需要更长的时间。

四、推广

在同一流动方向，如 y 向，还可将对流项与耗散项进行分裂计算，允许对流项用特征线法，而耗散项用其他方法。若物理方程中包含多种影响，如两相流中各相的动量方程、能量方程；燃烧流中的化学反应、物理反应均可用上述概念进行物理分裂计算，计算中各分裂式采用合适的格式，要求：

- 与原方程相容；
- 达到一定的精度；

· 稳定。

这就是物理分裂的概念。

3.3 二维抛物型方程的隐格式与 ADI 方法

3.3.1 隐格式与五对角稀疏矩阵

对于方程：

$$\frac{\partial U}{\partial t}=\overline{\gamma}_1\frac{\partial^2 U}{\partial x^2}+\overline{\gamma}_2\frac{\partial^2 U}{\partial y^2} \tag{3.21}$$

用 Crank-Nicolson 格式，即：

$$\exp\left[-\frac{1}{2}(D_x^2+D_y^2)\right]U_{j,m}^{n+1}=\exp\left[\frac{1}{2}(D_x^2+D_y^2)\right]U_{j,m}^n$$

若 $D_x^2=\delta_{xx}/\Delta x^2, D_y^2=\delta_{yy}/\Delta y^2$，则有：

$$\left[-\frac{r_x}{2}(U_{j+1,m}+U_{j-1,m})+(1+r_x+r_y)U_{j,m}-\frac{r_y}{2}(U_{j,m+1}+U_{j,m-1})\right]^{n+1}=q^n \tag{3.22}$$

其中 $r_x=\gamma_1\dfrac{\Delta t}{\Delta x^2}, r_y=\gamma_2\dfrac{\Delta t}{\Delta y^2}, q^n$ 由已知量 U^n 组成。(3.22)写成矩阵形式为：

$$\begin{pmatrix} \cdots & & & & & & & & & \\ 0-\frac{r_y}{2} & 0 & -\frac{r_x}{2} & 1+r_x+r_y & -\frac{r_x}{2} & 0 & -\frac{r_y}{2} & 0 & 0 \\ 0 & 0 & -\frac{r_y}{2} & 0 & -\frac{r_x}{2} & 1+r_x+r_y & -\frac{r_x}{2} & 0 & -\frac{r_y}{2} & 0 \\ \ddots & & & \ddots & & & & & & \\ & \ddots & & & \ddots & & & & & \\ & & \ddots & & & \ddots & & & & \\ & & & \ddots & & & \ddots & & & \\ & & & & \ddots & & & \ddots & & \\ & & & & & \ddots & & & \ddots & \end{pmatrix}\begin{pmatrix}\vdots \\ U_{j,m-1} \\ \vdots \\ U_{j-1,m} \\ U_{j,m} \\ U_{j+1,m} \\ \vdots \\ U_{j,m+1} \\ \vdots\end{pmatrix}^{n+1}=q^n$$

这是一个$(J-1)\times(M-1)$的大型五对角稀疏矩阵方程，计算比较困难，对有些必须用此方法求解的问题可直接求解，但对不少问题用较简便的 ADI 方法较好。

3.3.2 ADI 法(交替方向隐式法)

一、Peaceman-Rachford ADI(Alternating Direction Implicit)方法

它从 Crank-Nicolson 格式(3.22)出发：

$$\exp\left[-\frac{k}{2}(D_x^2+D_y^2)\right]U^{n+1}=\left[\exp\frac{k}{2}(D_x^2+D_y^2)\right]U^n$$

若 $D_x^2=\delta_{xx}/\Delta x^2, D_y^2=\delta_y/\Delta y^2$，则：

$$\left(1-\frac{r_x}{2}\delta_{xx}\right)\left(1-\frac{r_y}{2}\delta_{yy}\right)U^{n+1}=\left(1+\frac{r_x}{2}\delta_{xx}\right)\left(1+\frac{r_y}{2}\delta_{yy}\right)U^n \quad (3.23)$$

上式的截断误差首项的量级是 $O(k^2,h^2)$，具有时间和空间的二阶精度。略去$\frac{r_xr_y}{4}\delta_{xx}\delta_{yy}U^{n+1}$及$\frac{r_xr_y}{4}\delta_{xx}\delta_{yy}U^n$，因它们不影响(3.23)的精度，则有：

$$\left(1-\frac{r_x}{2}\delta_{xx}-\frac{r_y}{2}\delta_{yy}\right)U^{n+1}=\left(1+\frac{r_x}{2}\delta_{xx}+\frac{r_y}{2}\delta_{yy}\right)U^n$$

其中 $r_x=\gamma_1\frac{k}{h^2}$，$r_y=\gamma_2\frac{k}{h^2}$，$h=\Delta x=\Delta y$。

ADI 法为：

$$\left(1-\frac{r_x}{2}\delta_{xx}\right)U^{\overline{n+1}}=\left(1+\frac{r_y}{2}\delta_{yy}\right)U^n \quad (3.24a)$$

$$\left(1-\frac{r_y}{2}\delta_{yy}\right)U^{n+1}=\left(1+\frac{r_x}{2}\delta_{xx}\right)U^{\overline{n+1}} \quad (3.24b)$$

这样做每一步均解三对角矩阵。但必须指出，上述 ADI 法在前一步是条件稳定的，即 $r_y\leqslant1$。到第二步完成后才达到恒稳的要求。

二、Douglas-Rachford 法

用全隐式离散(3.21)，$[\exp-k(D_x^2+D_y^2)]U^{n+1}=U^n$。略去高阶项可写成：

$$(1-r_x\delta_{xx}-r_y\delta_{yy})U^{n+1}=U^n$$

进行 ADI 得：

$$(1-r_x\delta_{xx})U^{\overline{n+1}}=(1+r_y\delta_{yy})U^n \quad (3.25a)$$

$$(1-r_y\delta_{yy})U^{n+1}=U^{\overline{n+1}}-r_y\delta_{yy}U^n \quad (3.25b)$$

实质上右侧多了一项 $r_xr_y\delta_{xx}\delta_{yy}U^n$，其量级为 $O(h^4)$，略去不影响精度，当然保证恒稳。

格式(3.25)的每一步均是恒稳的，且可以引伸到三维问题中去，因此使用广泛。对三维流动：

$$(1-r_x\delta_{xx})U^{\overline{n+1}}=(1+r_y\delta_{xx}+r_z\delta_{zz})U^n \quad (3.26a)$$

$$(1-r_y\delta_{yy})U^{\overline{\overline{n+1}}}=U^{\overline{n+1}}-r_y\delta_{yy}U^n \quad (3.26b)$$

$$(1-r_z\delta_{zz})U^{n+1}=U^{\overline{\overline{n+1}}}-r_z\delta_{zz}U^n \quad (3.26c)$$

三、D'Yakonob 法

D'Yakonob 法与 Peaceman-Rachford ADI 方法的区别在于分裂方法不同，出发方程是(3.23)。它将(3.23)式分裂成：

$$\left(1-\frac{r_x}{2}\delta_{xx}\right)U^{\overline{n+1}}=\left(1+\frac{r_x}{2}\delta_{xx}\right)\left(1+\frac{r_y}{2}\delta_{yy}\right)U^n$$

$$\left(1-\frac{r_y}{2}\delta_{yy}\right)U^{n+1}=U^{\overline{n+1}}$$

3.3.3 ADI 法的边界处理

前面引入的(3.24)式简称 PR 法，(3.26)式简称 DR 法。两种方法的中间步与原方程不相容，这是很明显的，因为只离散了方程的一部分，因此 $U^{\overline{n+1}}$ 没有物理意义，只是计算

中的一个中间过程。但为了解出结果必须给出中间步的边界条件。如图 3.1，左边界 $g_0(x)$ 与右边界 $g_1(x)$ 均已知，对 DR 来说，$U_g^{\overline{n+1}}=(1-r_y\delta_{yy})U_g^{n+1}+r_y\delta_{yy}U_g^n$，在边界处 U_g^{n+1} 和 U_g^n 已知，故 $U_g^{\overline{n+1}}$ 可直接求出。对 PR，将 $U^{\overline{n+1}}$ 作未知数，使(3.24)式的二步相加得：

$$U_g^{\overline{n+1}}=\frac{1}{2}\Big[(1+\frac{r_y}{2}\delta_{yy})U_g^n+(1-\frac{r_y}{2}\delta_{yy})U_g^{n+1}\Big]$$

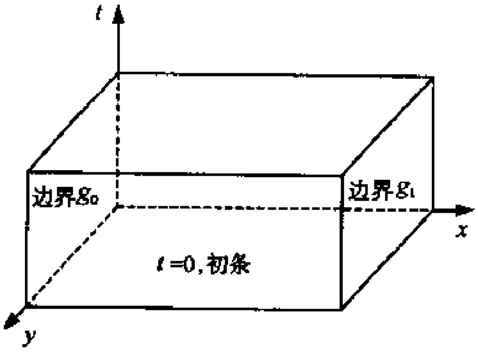


图 3.1

对定常问题，U_g 不随时间变化，则在边界处 $U_g^{\overline{n+1}}=U_g^n=U_g$。

3.4 二维双曲型与输运型方程的隐格式

流体力学中常见的非定常 Euler 方程是双曲型方程，N. S. 方程是输运型的。隐式求解方法很多，下面主要介绍近似因子法（Approximate Factor Method，简称 AF 法）和 MacCormack 隐格式等。

3.4.1 AF 法的基本原理

考虑方程：

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \tag{3.15}$$

用隐式离散，然后用 ADI 法求解。但是 F,G 是 U 的非线性函数，怎样将未知变量从 F^{n+1} 与 G^{n+1} 中分离出来而不严重影响其精度是一个大问题。Beam 等较好地处理了这一问题，现将其处理过程叙述如下。

一、由 Taylor 展开式写成隐式差分式

$$U^{n+1}=U^n+\left(\frac{\partial U}{\partial t}\right)^n\Delta t+\left(\frac{\partial^2 U}{\partial t^2}\right)\frac{\Delta t^2}{2}+O(\Delta t^3)$$

$$U^n=U[(n+1)\Delta t-\Delta t]=U^{n+1}-\left(\frac{\partial U}{\partial t}\right)^{n+1}\Delta t+\left(\frac{\partial^2 U}{\partial t^2}\right)^{n+1}\frac{\Delta t^2}{2}+O(\Delta t^3)$$

两式相减，得：

$$U^{n+1}-U^n=\frac{\Delta t}{2}\left[\left(\frac{\partial U}{\partial t}\right)^n+\left(\frac{\partial U}{\partial t}\right)^{n+1}\right]+\frac{\Delta t^2}{4}\left[\left(\frac{\partial^2 U}{\partial t^2}\right)^n-\left(\frac{\partial^2 U}{\partial t^2}\right)^{n+1}\right]+O(\Delta t^3)$$

近似取 $\left(\frac{\partial^2 U}{\partial t^2}\right)^n=\left(\frac{\partial^2 U}{\partial t^2}\right)^{n+1}$，由(3.15)得：

$$U^{n+1}=U^n-\frac{\Delta t}{2}\left[\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)^n+\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)^{n+1}\right]+O(\Delta t^3) \tag{3.27}$$

二、线化处理

上式中 $\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)^{n+1}$ 无法直接求解，处理如下：

$$F^{n+1}=F(U^{n+1})=F^n+\left(\frac{\partial F}{\partial t}\right)^n\Delta t+\left(\frac{\partial^2 F}{\partial t^2}\right)^n\frac{\Delta t^2}{2}$$

$$=F^n+A^n(U^{n+1}-U^n)+O(k^2) \tag{3.28a}$$

$$G^{n+1}=G^n+B^n(U^{n+1}-U^n)+O(k^2) \tag{3.28b}$$

其中 $A=\frac{\partial F}{\partial U},B=\frac{\partial G}{\partial U}$为 Jacobian 矩阵。

这是第一步线化，它将 U^{n+1}从 F^{n+1}与 G^{n+1}中取出，但时间上只有一阶精度。

三、F,G 为齐次函数情况下，即 $F(\lambda U)=\lambda F(U)$时的简化

$F=AU,G=BU$，代入(3.28)得：

$$F^{n+1}=F^n+A^n(U^{n+1}-U^n)=A^nU^{n+1},G^{n+1}=A^nU^{n+1} \tag{3.29}$$

再代入(3.27)得：

$$U^{n+1}=U^n-\frac{\Delta t}{2}\left[\left(\frac{\partial A\cdot}{\partial x}+\frac{\partial B\cdot}{\partial y}\right)^nU^n+\left(\frac{\partial A\cdot}{\partial x}+\frac{\partial B\cdot}{\partial y}\right)^nU^{n+1}\right]$$

即：

$$\left[I+\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}+\frac{\partial B\cdot}{\partial y}\right)^n\right]U^{n+1}=\left[I-\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}+\frac{\partial B\cdot}{\partial y}\right)^n\right]U^n$$

或写成：

$$\left[I+\frac{k}{2}\left(A^n\frac{\partial}{\partial x}+B^n\frac{\partial}{\partial y}\right)^n\right]U^{n+1}=\left[I-\frac{k}{2}\left(A^n\frac{\partial}{\partial x}+B^n\frac{\partial}{\partial y}\right)^n\right]U^n \tag{3.30}$$

四、ADI

在方程(3.30)中加上$\left(\frac{\partial A}{\partial x}\right)^n\left(\frac{\partial B}{\partial y}\right)^n\left(\frac{U^{n+1}-U^n}{\Delta t}\right)\frac{\Delta t^2}{4}=\left(\frac{\partial A\partial B\partial U}{\partial x\partial y\partial t}\right)^n\frac{\Delta t^2}{4}+O(\Delta t^3)$不影响原精度。关于 U^{n+1}的项加于左侧，关于 U^n 的项加于右侧，则由(3.30)可得：

$$\left[I+\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]\left[I+\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^{n+1}=\left[I-\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]\left[I-\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^n$$

再用 PR ADI 方法：

$$\left[I+\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]U^{\overline{n+1}}=\left[I-\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]U^n \tag{3.31a}$$

$$\left[I+\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^{n+1}=\left[I-\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^{\overline{n+1}} \tag{3.31b}$$

或 D'Yakonob 方法：

$$\left[I+\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]U^{\overline{n+1}}=\left[I-\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]\left[I-\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^n \tag{3.32a}$$

$$\left[I+\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^{n+1}=U^{\overline{n+1}} \tag{3.32b}$$

这是当时的 AF 格式，又称 Beam-Warming 格式，应用较广泛，但其中缺点较明显：

- 差分方程中守恒性作了近似的处理，不严格守恒；
- PR 分裂格式中间步与原方程不相容，使边界条件不好处理；
- 每步计算时间较长。

3.4.2 Delta 型 Beam-Warming 格式

为提高解题效率，保证中间步相容，要对格式作些修改。

先证明(3.32)第一式$\left[I+\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]U^{\overline{n+1}}=\left[I-\frac{k}{2}\left(\frac{\partial A\cdot}{\partial x}\right)^n\right]\left[I-\frac{k}{2}\left(\frac{\partial B\cdot}{\partial y}\right)^n\right]U^n$与原方程不相容。用 Taylor 级数展开上式左右两侧得：

$$U^n+k\frac{\partial U}{\partial t}+\frac{k}{2}\frac{\partial A^n}{\partial x}\left(U^n+k\frac{\partial U}{\partial t}\right)+\cdots$$

$$=U^n-\frac{k}{2}\left(\frac{\partial A^n}{\partial x}\cdot U^n+\frac{\partial B^n}{\partial y}\cdot U^n\right)+\frac{k^2}{4}\left(\frac{\partial A^n\cdot}{\partial x}\frac{\partial B^n\cdot}{\partial y}\right)U^n+\cdots$$

令 $k\to 0$ 得：$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{1}{2}\frac{\partial G}{\partial y}=0$，故与原方程(3.15)不相容。

1978 年 Beam 等对(3.31)与(3.32)提出如下修改：

令 $\delta U^n=U^{n+1}-U^n$，将(3.27)改写为：

$$\delta U^n=-\frac{k}{2}\left[2\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)^n+\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)^{n+1}-\left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)^n\right]$$

令 $\delta F^n=F^{n+1}-F^n=A^n(U^{n+1}-U^n)=A^n\delta U^n$，$\delta G^n=B^n\delta U^n$，得

$$\delta U^n=-\frac{k}{2}\left[2\left(\frac{\partial A^n\cdot}{\partial x}+\frac{\partial B^n\cdot}{\partial y}\right)U^n+\left(\frac{\partial A^n\cdot}{\partial x}+\frac{\partial B^n\cdot}{\partial y}\right)\delta U^n\right]$$

所以，

$$\left[I+\frac{k}{2}\left(\frac{\partial A^n\cdot}{\partial x}+\frac{\partial B^n\cdot}{\partial y}\right)\right]\delta U^n=-k\left(\frac{\partial A^n\cdot}{\partial x}+\frac{\partial B^n\cdot}{\partial y}\right)U^n \tag{3.33}$$

按 ADI 方法：

$$\left(I+\frac{k}{2}\frac{\partial A^n\cdot}{\partial x}\right)\delta U^{\bar{n}}=-k\left(\frac{\partial A^n\cdot}{\partial x}+\frac{\partial B^n\cdot}{\partial y}\right)U^n \tag{3.34a}$$

$$\left(I+\frac{k}{2}\frac{\partial B^n\cdot}{\partial y}\right)\delta U^n=\delta U^{\bar{n}} \tag{3.34b}$$

$$U^{n+1}=U^n+\delta U^n \tag{3.34c}$$

这称为 Delta 型 Beam-Warming 格式。可证明中间步与原方程相容，如果对$\frac{\partial}{\partial x}$和$\frac{\partial}{\partial y}$采用中心差分的形式，则这就是解两个块三对角矩阵的问题。A，B 中块的大小由 U 中所包括的变量数目决定。如对一般二维 Euler 或 N.S. 方程，A 和 B 是 4×4 矩阵。

3.4.3 Beam-Warming 格式的差分方程

用原来的 Beam-Warming 隐格式求解问题只能保证时间一阶精度，其空间精度与空间差分格式有关。为保证空间二阶精度，取中心差分，则其差分方程(3.32)形式如下：

$$U^{\overline{n+1}}=\left[I-\frac{k}{2}\frac{\partial B^n\cdot}{\partial y}\right]U^n$$

$$\left[I+\frac{k}{2}\frac{\partial A^n\cdot}{\partial x}\right]U^{\overline{\overline{n+1}}}=\left[I-\frac{k}{2}\frac{\partial A^n\cdot}{\partial x}\right]U^{\overline{n+1}}$$

$$\left[I+\frac{k}{2}\frac{\partial B^n\cdot}{\partial y}\right]U^{n+1}=U^{\overline{\overline{n+1}}}$$

即：

$$U_{j,m}^{\overline{n+1}}=U_{j,m}^n-\frac{k}{4\Delta y}\left(B_{j,m+1}^nU_{j,m+1}^n-B_{j,m-1}^nU_{j,m-1}^n\right) \tag{3.35a}$$

$$U^{\overline{\overline{n+1}}}_{j,m}+\frac{k}{4\Delta x}\left(A^n_{j+1,m}U^{\overline{\overline{n+1}}}_{j+1,m}-A^n_{j-1,m}U^{\overline{\overline{n+1}}}_{j-1,m}\right) \tag{3.35b}$$

$$=U^{\overline{n+1}}_{j,m}-\frac{k}{4\Delta x}\left(A^n_{j+1,m}U^{\overline{n+1}}_{j+1,m}-A^n_{j-1,m}U^{\overline{n+1}}_{j-1,m}\right)$$

$$U^{n+1}_{j,m}+\frac{k}{4\Delta y}\left(B^n_{j,m+1}U^{n+1}_{j,m+1}-B^n_{j,m-1}U^{n+1}_{j,m-1}\right)=U^{\overline{\overline{n+1}}}_{j,m} \tag{3.35c}$$

可以证明，此隐格式是无条件稳定的。计算表明，在物理量变化剧烈的地方，无论是原格式(3.31)、(3.32)，还是1978年提出的 Delta 格式(3.34)，每项后面均要加4阶人工耗散项。

通常加于(3.35)右侧的值分别为：

(3.35a)：$-\frac{w_y}{16}\delta_y^4U^n_{j,m}=-\frac{w_y}{16}\left[U^n_{j,m+2}-4U^n_{j,m+1}+6U^n_{j,m}-4U^n_{j,m-1}+U^n_{j,m-2}\right]$

(3.35b)：$-\frac{w_x}{8}\delta_x^4U^{\overline{n+1}}_{j,m}$

(3.35c)：$-\frac{w_y}{8}\delta_y^4U^{\overline{\overline{n+1}}}_{j,m}$

加负值是为了保证不致失稳。w_x，w_y 是常数，通常是小于1的正数。

3.4.4 广义 ADI 法——三维流动的 Delta 格式

三维流的 Beam-Warming 格式的一般形式是

$$(I+\Lambda_x+\Lambda_y+\Lambda_z)\left(U^{n+1}_{j,m,l}-U^n_{j,m,l}\right)=RU^n_{j,m,l}$$

式中 $\Lambda_x=\frac{k}{2}\frac{\partial A^n\cdot}{\partial x}$，$\Lambda_y=\frac{k}{2}\frac{\partial B^n\cdot}{\partial y}$，$\Lambda_z=\frac{k}{2}\frac{\partial C^n\cdot}{\partial z}$，$R$ 是包括 $\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}$ 在内的综合算子。

广义的 ADI 分裂：

$$(I+\Lambda_x)\delta U^{**}_{j,m,l}=RU^n_{j,m,l} \tag{3.36a}$$

$$(I+\Lambda_y)\delta U^{*}_{j,m,l}=\delta U^{**}_{j,m,l} \tag{3.36b}$$

$$(I+\Lambda_z)\delta U^{n}_{j,m,l}=\delta U^{*}_{j,m,l} \tag{3.36c}$$

$$U^{n+1}_{j,m,l}=\delta U^n_{j,m,l}+U^n_{j,m,l} \tag{3.36d}$$

3.4.5 Briley-McDonald 隐格式

此格式与 Beam-Warming 格式类似，但上述 Beam 格式看上去犹如 Crank-Nicolson 型格式，即有一半是未知隐式差分，另有一半是显式差分；Briley 格式基本上是全隐式，因此，它的 ADI 更接近于 Douglas-Rachford 分裂。下面以 N.S. 方程(1.1)$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0$ 为例，介绍其形式及求解方法。

一、Briley 方法的差分

主要采用全隐方法。下面以质量方程和流向动量方程为例进行说明。

$$\frac{1}{k}(\rho^{n+1}-\rho^n)=-\frac{\partial}{\partial x}(\rho u)^{n+1}-\frac{\partial}{\partial y}(\rho v)^{n+1}$$

$$\frac{1}{k}[(\rho u)^{n+1}-(\rho u)^{n}]=-\frac{\partial}{\partial x}\left[p+\rho u^{2}-\mu\left(\frac{4}{3}\frac{\partial u}{\partial x}-\frac{2}{3}\frac{\partial v}{\partial y}\right)\right]^{n+1}$$
$$-\frac{\partial}{\partial y}\left[\rho uv-\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right]^{n+1}$$

……

二、线化

上式中未知量以非线性形式组合，不好求解。用 Taylor 展开，如：

$$(\rho u)^{n+1}=(\rho u)^{n}+\left(\frac{\partial \rho u}{\partial t}\right)^{n}\cdot k+\cdots$$
$$=(\rho u)^{n}+\left(\rho\frac{\partial u}{\partial t}+u\frac{\partial \rho}{\partial t}\right)^{n}\cdot k+\cdots$$
$$=(\rho u)^{n}+\left(\rho^{n}\cdot\frac{u^{n+1}-u^{n}}{k}+u^{n}\cdot\frac{\rho^{n+1}-\rho^{n}}{k}\right)$$
$$=\rho^{n}u^{n+1}+\rho^{n+1}u^{n}-\rho^{n}u^{n}$$

这里凡带"n"的均为已知值，凡带"$n+1$"的均为未知值。其他组合变量均以此类推进行线化，最后得（带"n"的不再写出"n"）：

$$\rho^{n+1}-\rho=k\left[-\frac{\partial}{\partial x}(\rho u^{n+1}+u\rho^{n+1}-\rho u)-\frac{\partial}{\partial y}(\rho v^{n+1}+v\rho^{n+1}-\rho v)\right] \quad (3.37a)$$

$$\rho^{n+1}u-u^{n+1}\rho-2\rho u=k\left[-\frac{\partial}{\partial x}(\rho^{n+1}u^{2}+2\rho uu^{n+1}-2\rho u^{2}+p^{n+1}-\tau_{xx}^{n+1})\right]$$
$$+k\left[-\frac{\partial}{\partial y}(\rho^{n+1}uv+\rho vu^{n+1}+\rho uv^{n+1}-2\rho uv-\tau_{xy}^{n+1})\right] \quad (3.37b)$$

……

三、ADI

归纳起来具有$(I+\Lambda_x+\Lambda_y)\bar{U}^{n+1}=H\bar{U}^{n}$ 的形式。H 只是一般简单的差分，通常不包括矩阵运算。$\bar{U}=(\rho,u,v,p)^{T}$，因此其分裂可写成(3.25)式的形式：

$$(I+\Lambda_x)\bar{U}^{\overline{n+1}}=(I+\Lambda_y)\bar{U}^{n}$$
$$(I+\Lambda_y)\bar{U}^{n+1}=\bar{U}^{\overline{n+1}}-\Lambda_y\bar{U}^{n}$$

对质量方程的 ADI 式是：

$$\rho^{\overline{n+1}}-\rho=k[-\delta_x(\rho u^{\overline{n+1}}+u\rho^{\overline{n+1}}-\rho u)+\delta_x(-\rho v)]$$
$$\rho^{n+1}-\rho^{\overline{n+1}}=k[-\delta_y(\rho v^{n+1}+v\rho^{n+1}-\rho v)+\delta_y(-\rho v)]$$

对动量方程：

$$\frac{1}{k}(\rho u^{\overline{n+1}}+u\rho^{\overline{n+1}}-2\rho u)=-\delta_x[\rho^{n+1}u^{2}+2\rho uu^{n+1}-2\rho u^{2}+p^{\overline{n+1}}-\tau_{xx}^{*}]$$
$$+\delta_y[-\tau_{xy}^{*}-2\rho uv]$$

$$\frac{1}{k}(\rho u^{n+1}+u\rho^{n+1}-\rho u^{\overline{n+1}}-u\rho^{\overline{n+1}})=\delta_y[-\tau_{xy}^{*}-2\rho uv]$$
$$-\delta_y[\rho^{n+1}uv+u^{n+1}\rho v+v^{n+1}\rho u]$$

在这里，"τ"没有作分裂变量处理。在 τ_{xx}^{*}中仅对$\frac{\partial u}{\partial x}$项作了$\overline{n+1}$处理，$\frac{\partial v}{\partial x}$项用"$n$"的值；

τ_{xy}^*中$\frac{\partial}{\partial y}$项用"$n+1$"层的值。

四、隐式差分方程的求解

上述差分方程可改写，δ_x，δ_y 分别用中心差分代之。写出差分方程后可得：

$$a_{j,m}^n \overline{\phi_{j+1,m}^{n+1}} + b_{j,m}^n \overline{\phi_{j,m}^{n+1}} + c_{j,m}^n \overline{\phi_{j-1,m}^{n+1}} = d_{j,m}^n \tag{3.38a}$$

$$\alpha_{j,m}^n \phi_{j,m+1}^{n+1} + \beta_{j,m}^n \phi_{j,m}^{n+1} + \gamma_{j,m}^n \phi_{j,m-1}^{n+1} = \overline{\xi_{j,m}^{n+1}} + \delta_{j,m}^n \tag{3.38b}$$

式中 $\phi=(\rho,u,v,e)^T$，$a,b,c,\alpha,\beta,\gamma$ 均由 n 层变量的值所组成，为已知值。$\overline{\xi_{j,m}^{n+1}}$为(3.38a)式求得的值。对二维可压流问题，$a,b,c,\alpha,\beta,\gamma$ 均是 4×4 矩阵。

求解第一式相当于求解下列三对角块矩阵方程：

$$\begin{bmatrix} b_1 & a_1 & 0 & 0 \\ c_2 & b_2 & a_2 & 0 \\ 0 & c_3 & b_3 & a_3 \\ 0 & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \overline{\phi_{1,m}^{n+1}} \\ \overline{\phi_{2,m}^{n+1}} \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} d_{1,m}^n \\ d_{2,m}^n \\ \vdots \\ \vdots \end{bmatrix}$$

求解可以用追赶法，但对这类问题也可直接用三角分解法求解，即：

$$Ax = b \Rightarrow LUx = b$$

其中 L 是下三角矩阵，U 是上三角矩阵：

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & 0 \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & u_{n-1n} \\ 0 & \cdots & \cdots & 0 & u_{nn} \end{bmatrix}$$

只要 A 是非奇异矩阵，各元素 $a_{i,j}$已知，则各 $l_{i,j}$和 $u_{i,j}$均可用 Crout LU 分解求出。于是上述问题变成：

$$A\overline{\phi^{n+1}} = d^n, \quad LU\overline{\phi^{n+1}} = d^n$$

令 $\overline{y^{n+1}}=U\overline{\phi^{n+1}}$，则：$L\overline{y^{n+1}}=d^n$，解出 $\overline{y^{n+1}}$。再由 $U\overline{\phi^{n+1}}=\overline{y^{n+1}}$解出 $\overline{\phi^{n+1}}$。

上述计算过程都有标准程序可用，而方法则可在线性代数与数值方法的教科书中找到。

3.4.6 MacCormack 隐格式

仍考虑

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0 \tag{1.1}$$

$$\frac{\partial}{\partial t}U = -\left(\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y}\right) = -\left(\frac{\partial A\cdot}{\partial x} + \frac{\partial B\cdot}{\partial y}\right)U$$

$$\frac{\partial}{\partial t}\left(\frac{\partial U}{\partial t}\right) = -\left(\frac{\partial A\cdot}{\partial x} + \frac{\partial B\cdot}{\partial y}\right)\left(\frac{\partial U}{\partial t}\right)$$

用隐格式解上述方程：

$$\left(\frac{\partial U}{\partial t}\right)^{n+1} - \left(\frac{\partial U}{\partial t}\right)^{n} = -k\left(\frac{\partial A\cdot}{\partial x} + \frac{\partial B\cdot}{\partial y}\right)\left(\frac{\partial U}{\partial t}\right)^{n+1}$$

$$\left[I + k\left(\frac{\partial A\cdot}{\partial x} + \frac{\partial B\cdot}{\partial y}\right)\right]\left(\frac{\partial U}{\partial t}\right)^{n+1} = \left(\frac{\partial U}{\partial t}\right)^{n}$$

令 $\delta U^n = k\left(\frac{\partial U}{\partial t}\right)^n = k\ \frac{U^{n+1}-U^n}{k} = U^{n+1} - U^n$，$\delta U^{n+1} = k\left(\frac{\partial U}{\partial t}\right)^{n+1}$，则 $\left[I+k\left(\frac{\partial A\cdot}{\partial x}+\frac{\partial B\cdot}{\partial y}\right)\right]\delta U^{n+1}=\delta U^n$。在不影响精度的条件下可写成：

$$\left[I + k\frac{\partial A\cdot}{\partial x}\right]\left[I + k\frac{\partial B\cdot}{\partial y}\right]\delta U^{n+1} = \delta U^n \tag{3.39}$$

然后用 MacCormack 格式求解。先写显式：

$$\delta U^n = U^{n+1} - U^n = -k\left(\frac{\partial F^n}{\partial x} + \frac{\partial G^n}{\partial y}\right)$$

$$U^{\overline{n+1}} = U^n - k\left(\frac{\triangle_x F^n}{h_x} + \frac{\triangle_y G^n}{h_y}\right)$$

代入(3.39)得：

预测步：

$$\begin{cases}\delta U^n = -k\left(\dfrac{\triangle_x F^n}{h_x} + \dfrac{\triangle_y G^n}{h_y}\right)\\ \left(I - k\dfrac{\triangle_x|A|^n}{h_x}\right)\left(I - k\dfrac{\triangle_y|B|^n}{h_y}\right)\delta U^{\overline{n+1}} = \delta U^n\\ U^{\overline{n+1}} = U^n + \delta U^{\overline{n+1}}\end{cases}$$

校正步：

$$\begin{cases}\delta U^{\overline{n+1}} = -k\left(\dfrac{\nabla_x F^{\overline{n+1}}}{h_x} + \dfrac{\nabla_y G^{\overline{n+1}}}{h_y}\right)\\ \left(I + k\dfrac{\nabla_x|A|^n}{h_x}\right)\left(I + k\dfrac{\nabla_y|B|^n}{h_y}\right)\delta U^{n+1} = \delta U^{\overline{n+1}}\\ U^{n+1} = \dfrac{1}{2}(U^n + U^{\overline{n+1}} + \delta U^{n+1})\end{cases} \tag{3.40}$$

这里在 A 与 B 上加 $|A|$，$|B|$，表示 A 与 B 先与括号外的 δU^{n+1} 等相乘再取向前或向后差分。

此格式的优点：隐式是一侧差分，不必用系数矩阵求逆求解，效率较高；缺点：要事先进行大量矩阵运算，边界条件处理困难。因此使用并不广泛。

3.5 常用其他形式格式

3.5.1 不相容于非定常流的格式

为了说明问题，先用一维流$\frac{\partial u}{\partial t}+\frac{\partial}{\partial x}\left(\frac{u^2}{2}\right)=\bar{\gamma}\frac{\partial^2 u}{\partial x^2}$来求解。对时间相关法而言，是为解决定常流方程类型不确定问题而加上$\frac{\partial u}{\partial t}$的，因此有些格式在定常时相容于原方程，非定常时不相容也可采用。如：

$$u_m^{n+1}=u_m^n-\frac{\alpha}{2}\left[\left(\frac{u^2}{2}\right)_{m+1}^{n+1}-\left(\frac{u^2}{2}\right)_{m-1}^{n+1}\right]+r\left[u_{m+1}^n-2u_m^n+u_{m-1}^n\right]$$

其中$\alpha=\frac{\Delta t}{\Delta x}$，$r=\bar{\gamma}\frac{\Delta t}{\Delta x^2}$。Taylor 展开并令 $\Delta t\to 0$，$\Delta x\to 0$ 得：

$$\frac{\partial u}{\partial t}+K\left[\frac{\partial}{\partial x}\left(\frac{u^2}{2}\right)-\bar{\gamma}\frac{\partial^2 u}{\partial x^2}\right]=0,K=\frac{1}{1-(r+u\alpha/2)}$$

由此可见，非定常时与原方程不相容，定常时与原方程相容。但用这种格式解人为非定常流的定常问题还是有用的。

对于二维流动，方程为：

$$\frac{\partial u}{\partial t}+\frac{\partial}{\partial x}\left(\frac{u^2}{2}\right)+\frac{\partial}{\partial y}\left(\frac{u^2}{2}\right)=\bar{\gamma}_1\frac{\partial^2 u}{\partial x^2}+\bar{\gamma}_2\frac{\partial^2 u}{\partial y^2} \tag{3.41}$$

对应的格式为：

$$u_{j,m}^{n+1}=u_{j,m}^n-\left\{\frac{\alpha_1}{2}\left[\left(\frac{u^2}{2}\right)_{j+1,m}^{n+1}-\left(\frac{u^2}{2}\right)_{j-1,m}^{n+1}\right]+\frac{\alpha_2}{2}\left[\left(\frac{u^2}{2}\right)_{j,m+1}^{n+1}-\left(\frac{u^2}{2}\right)_{j,m-1}^{n+1}\right]\right\}$$
$$+r_1(u_{j+1,m}^n-2u_{j,m}^n+u_{j-1,m}^n)+r_2(u_{j,m+1}^n-2u_{j,m}^n+u_{j,m-1}^n) \tag{3.42}$$

其中$\alpha_1=\frac{\Delta t}{\Delta x}$，$\alpha_2=\frac{\Delta t}{\Delta y}$，$r_1=\bar{\gamma}_1\frac{\Delta t}{\Delta x^2}$，$r_2=\bar{\gamma}_2\frac{\Delta t}{\Delta y^2}$。

3.5.2 半隐 MacCormack 格式

对上节的一维流方程也可采用半隐的 MacCormack 格式：

预测：$$\overline{u_m^{n+1}}=u_m^n-\frac{\alpha}{2}\left({u_{m+1}^n}^2-{u_m^n}^2\right)+r\left(u_{m+1}^n-2\overline{u_m^{n+1}}+u_{m-1}^n\right)$$

校正：$$u_m^{n+1}=\frac{1}{2}\left[u_m^n+\overline{u_m^{n+1}}-\frac{\alpha}{2}\left(\overline{u_m^{n+1}}^2-\overline{u_{m-1}^{n+1}}^2\right)+r\left(\overline{u_{m+1}^{n+1}}-2u_m^{n+1}+\overline{u_{m-1}^{n+1}}\right)\right] \tag{3.43}$$

只在耗散项中的一项用了隐式，放宽了稳定性限制，但求解时仍可用类似于显式的解，即：

预测：$$\overline{u_m^{n+1}}=\frac{1}{1+2r}\left[u_m^n-\frac{\alpha}{2}\left({u_{m+1}^n}^2-{u_m^n}^2\right)+r\left(u_{m+1}^n+u_{m-1}^n\right)\right]$$

校正步也可类似求解。这个格式得到的差分方程只在定常时才与原方程相容。同理，对于二维流动的(3.41)式也可用类似的方法求解，这里不再罗列。

由于前面讲的 MacCormack 隐格式也是用类似于显式的直接求解，因此也有人将这

种格式统称为半隐式。

3.6 网格设计与边界条件

网格设计与边界条件是有限差分方法应用中的重要问题，其中边界条件在这里只能介绍简要的原则，有许多问题是在实际计算中不断总结与改进的；网格设计更是解决实际问题首先应处理好的技术问题，这里只介绍一般的概念，后面陆续介绍其新发展。

3.6.1 网格设计

用差分方程求解时，解域的网格取直线，边界也取直线，但实际上物面均是不规则曲面形的，从而存在如何处理边界，将网格划分与边界处理结合在一起的问题。

一、光滑流场的插值计算

1. Dirichlet 问题

当曲线边界 L 上各点物理量已知，求其邻点的物理量。如图 3.2 所示，解区中 H 点的值就由线性插值 $\phi_H=\frac{x_H-x_G}{x_{H1}-x_H}\phi_{H1}+\frac{x_H-x_{H1}}{x_G-x_{H1}}\phi_G$ 求得。其导数 $\left(\frac{\partial\phi}{\partial x}\right)_H=\frac{\phi_{H1}-\phi_H}{|HH_1|}$。

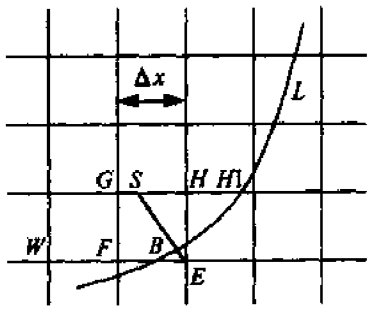


图 3.2

2. Neumann 问题

边界上法向导数已知，求邻近内点的各向导数。过边界外 E 点（邻近）作 L 的垂线交 L 于 B，交内边界线于 S。$\overline{HS}=a\Delta x$ 已知。

$$\phi_S=a\phi_G+(1-a)\phi_H$$

$$\left.\frac{\partial\phi}{\partial n}\right|_B=C_B=\frac{(\phi_E-\phi_S)}{\sqrt{a^2\Delta x^2+\Delta y^2}}=\frac{(\phi_E-\phi_S)}{l}$$

所以 $\phi_E=C_B\cdot l+\phi_S=C_B\cdot l+a\phi_G+(1-a)\phi_H$，可求出 ϕ_E。

$$\left.\frac{\partial\phi}{\partial x}\right|_F=\frac{\phi_E-\phi_W}{2\Delta x}=\frac{C_Bl+a\phi_G+(1-a)\phi_H-\phi_W}{2\Delta x}$$

也可以用差分格式迭代的办法，但这些办法只能用在物理量 ϕ 在边界上变化均匀的情况下。如果变化复杂、剧烈，那么用不同的导数取法或选取不同的取值点时，结果差异极大，同一边界条件甚至会得出不同的解，因此上述方法目前已较少使用了，代之以坐标变换方法。

二、几何关系的坐标变换

将物理面上的曲线形边界所围的解区，由 (x,y) 用确定几何关系式转变为计算平面上的矩形解区，如图 3.3 中物理面上的尾喷流流场，其控制方程为 Euler 方程组(3.15)：

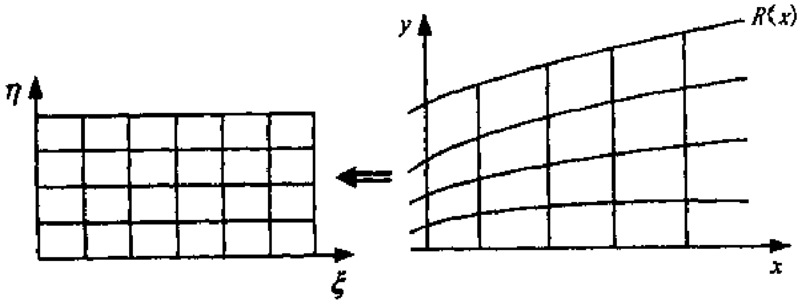
$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \tag{3.15}$$

设 $\xi=\xi(x),\eta=\eta(y)$。对图示问题 $\xi\equiv x,\eta=y/R(x),\tau\equiv t$，$R$ 是上边界距轴线 x 的间

距。$\xi_x=1,\eta_x=-\frac{R'}{R^2}y,\xi_y=0,\eta_y=\frac{1}{R}$，所以，$F_x=F_\xi\xi_x+F_\eta\eta_x=F_\xi-\frac{R'}{R^2}yF_\eta,G_y=G_\xi\xi_y+F_\eta\eta_y=\frac{1}{R}F_\eta$，代入原方程得：

$$\frac{\partial U}{\partial \tau}+\frac{\partial F}{\partial \xi}-\frac{R'}{R^2}y\frac{\partial F}{\partial \eta}+\frac{1}{R}\frac{\partial G}{\partial \eta}=0 \tag{3.44}$$

然后在(ξ,η)计算面上进行求解。这样，计算面上是矩形网格，求解合理。对不同形状的边界可采用不同的$\xi(x),\eta(y)$，但在很多复杂问题中，$\xi(x),\eta(y)$的形式不易确定。



(a) 计算面　　　　(b) 物理面

图 3.3

三、保角变换法

在(x,y)笛卡尔坐标面内有两族曲线$\phi(x,y)=C_1,\psi(x,y)=C_2$，二者互相正交，且满足 Cauchy-Rieman 条件，构成了复位函数：

$$\omega(z)=\phi(x,y)+i\psi(x,y)$$

用转换关系$z=f(\zeta)$将z平面内的两族曲线转换到ζ平面内的两族正交线：

$$\omega(f(\zeta))=\omega(\zeta)=\Phi(\xi,\eta)+i\Psi(\xi,\eta)$$

此时$\Phi(\xi,\eta)=C_1$与$\Psi(\xi,\eta)=C_2$形成的曲线与$\phi(x,y)=C_1,\psi(x,y)=C_2$两族线形状不同。如找到的合适的$z=f(\zeta)$使$\zeta$平面上矩形解域转绘到$z$平面上要求的形状，即达到目的，使计算在矩形网格下进行。

例　定常二维可压无旋流流过管壁曲线为双曲线的管道，如何将解区网格转换为矩形解区。管壁曲线的方程为：

$$y_w^2-x_w^2\tan^2 0.6=\sin^2 0.6 \tag{3.45}$$

设z是物理平面，其解区内有复位函数$\omega(z)=\phi(x,y)+i\psi(x,y)$存在，由$z=\mathrm{sh}\zeta$将其转给到计算平面，即：

$$z=x+iy=\mathrm{sh}(\xi+i\eta),$$
$$x=\mathrm{sh}\xi\cos\eta,y=\mathrm{ch}\xi\sin\eta \tag{3.46}$$

这就是计算平面上各点(ξ,η)与物理平面上各点(x,y)的对应关系。

$$\omega(z)=\omega(\mathrm{sh}\zeta)=\Phi(\xi,\eta)+i\Psi(\xi,\eta)$$

取(ξ,η)平面上的矩形网格区，即：$\Phi(\xi,\eta)=C_1=\xi$与$\Psi(\xi,\eta)=C_2=\eta$。由$\omega(\zeta)=\omega(z)$得：$\Phi=\phi=\xi=C_1,\Psi=\psi=\eta=C_2$。当$\xi=C_1$时，代入(3.46)得：

$$\frac{x^2}{\mathrm{sh}^2\xi}+\frac{y^2}{\mathrm{ch}^2\xi}=1$$

当 $\eta=C_2$ 时，代入(3.46)得：

$$-\frac{x^2}{\cos^2\eta}+\frac{y^2}{\sin^2\eta}=1$$

可以证明，当 $\eta=0.6$ 时，此方程即(3.45)。

两种平面的对应关系如图 3.5 所示。然后将控制方程按 $z=\mathrm{sh}\zeta$ 的要求进行变换、计算。

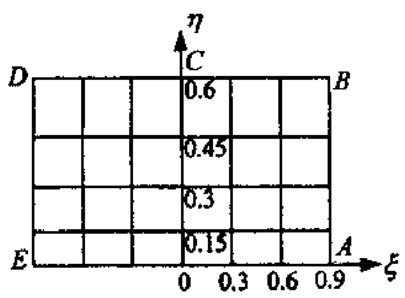


图 3.4　ζ 平面上矩形网格

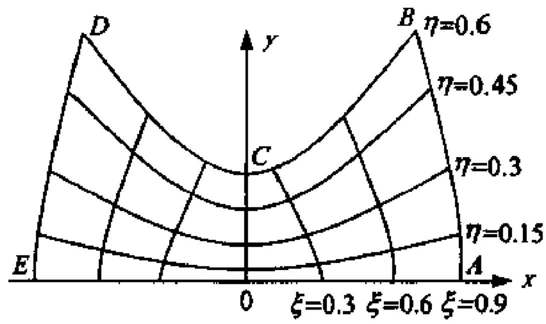


图 3.5　z 平面上网格

四、Thompson 边界拟合法

上述保角变换法是由变换关系式 $z=f(\zeta)$ 将计算平面的 $\xi=C_1,\eta=C_2$ 矩形域转到物理面(x,y)上要求的几何形状的解域。关键是，对较复杂的形状，$z=f(\zeta)$不易求得，且不能应用到三维去。

我们进行坐标转换的主要目的是保证物理面上物体边界域正好与曲线网格线相重合，这样边界的处理就不会有误差。如气流流过一个二维圆柱体的问题(见图 3.6)，最好应用极坐标(r,θ)，它保证在 $r=C_1$ 下是物理边界，那么进行边界处理时误差引入就小了。解决绕球体流动时则要用球坐标。

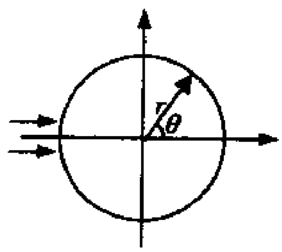


图 3.6

现假设气流流过一任意形状的边界，其 $\xi=\xi(x,y),\eta=\eta(x,y)$应如何选择呢？我们首先保证边界线与坐标曲线一致。设 $\eta=\eta_0=C_1$ 是物面边界，$\eta=\eta_\infty=C_2$ 是外边界。$\eta_0\leqslant\eta\leqslant\eta_\infty$ 布满了各坐标线。那么，封闭区 $\eta_0<\eta<\eta_\infty,\xi_0<\xi<\xi_\infty$ 内各坐标点应如何决定？它们必须完全受物面边界与外边界的控制。如椭圆型偏微分方程，它的内域完全由边界值所控制。最简单的椭圆型方程是 Laplace 方程，求解 $\xi=\xi(x,y),\eta=\eta(x,y)$的关系即转变为解：

$$\frac{\partial^2\xi}{\partial x^2}+\frac{\partial^2\xi}{\partial y^2}=0 \text{ 与 } \frac{\partial^2\eta}{\partial x^2}+\frac{\partial^2\eta}{\partial y^2}=0$$

在内边界(物面)(见图 3.7)已知：$\eta=\eta_0=C_1,\xi=f_0(x,y)$

在外边界已知：$\eta=\eta_\infty=C_2,\xi=f_\infty(x,y)$

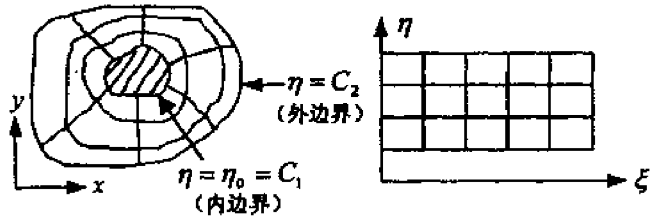


图 3.7

它能满足下列要求：正确地表达边界形状，网格正交，转换后控制方程形式相同，并且改用 Poisson 方程后可控制网格密度，可用差分方法进行数值计算，光滑物面网格能贴体。这方面的发展，以后再叙述。

3.6.2 边界条件

流体力学边界分两大类，一是渗透性边界，如来流、出流、激波等，允许在边界处有质量交换；另一种是非渗透性边界，如自由表面、对称面、面壁面等。现以固壁面为例说明边界条件的处理。

通常进行数值计算时，物理量 p,ρ,T,u,v 并非只能布置在网格点上或格心上，也可以将压力等标量分布在格心上，矢量 u,v 分布在格子边上，这种分布办法称为非一致性配置；另一种则物理量全配置在一点处，称为一致性配置。其道理以后再讲。下面以一致性配置为例。

一、无粘流直壁边界条件的提法

$(v_n)_w=0,(v_n)_w$ 表示固壁面上的法向分速。

上述条件常用镜像法确定，保证边界是条流线。在二维情况下，j 为 x 向注标，m 为 y 向注标。设 m_0 处为直壁面（见图 3.8），则：

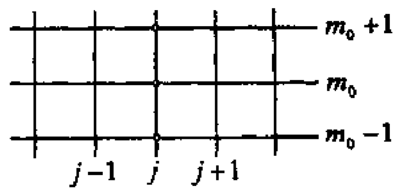


图 3.8

$$v_{j,m_0+1}=-v_{j,m_0-1}(v_{j,m_0}\equiv 0)$$

$$u_{j,m_0+1}=u_{j,m_0-1}\left(\frac{\partial u}{\partial y}\bigg|_w\equiv 0\right)$$

$$p_{j,m_0+1}=p_{j,m_0-1}\left(-\frac{\partial p}{\partial y}\bigg|_w=\left(v\frac{\partial u}{\partial y}+u\frac{\partial v}{\partial x}\right)\bigg|_w=0\right)$$

对曲线壁面应当用 $p_{j,m_0}=u^2_{j,m_0}/r$，r 为曲率半径。

ρ,T 的处理要小心，通常可按 $\frac{\partial\rho}{\partial y}\Big|_w=0$ 来达到。(j,m_0-1) 处的值是虚拟的。

二、粘流直壁边界条件

对连续介质流动，当物面静止时，

① 速度：$v_w=0$，其中 v_w 是物面当地速度矢；

② 温度：绝热壁 $T=T_w$；有热流：$q_w=-\lambda\left(\frac{\partial T}{\partial n}\right)_w$；

③ 密度：$\rho_w=3\rho_{j,m_0+1}-3\rho_{j,m_0+2}+\rho_{j,m_0+3}$，即 $\left(\frac{\partial^3\rho}{\partial y^3}\right)_w\equiv 0$；

④ 压强：大雷诺数边界层内 $p_w=p_{j,m_0+2}$，即$\left.\frac{\partial p}{\partial n}\right|_w=0$。

注意：边界点与内点格式结合后要考虑边界上的稳定性。其实，边界条件的处理有解析边界条件和数值边界条件之分，对不同的问题有多种处理技术，这里只作了最简要的说明。即这里所介绍的边界条件处理是可供选择的处理方法之一，其可行性要由数值计算的结果来验证。

三、二维超声速外边界的近似特征线法

在超声速流中，外边界通常有特征线存在。这里介绍一种工程上可用的处理上边界和外边界的方法。参见图 3.9，$\phi_{j-1,m}$，$\phi_{j-1,m-1}$，$\phi_{j,m-1}$已知，如何定 $\phi_{j,m}$，这里 ϕ 可代表 ρ,p,u,v 等物理量。

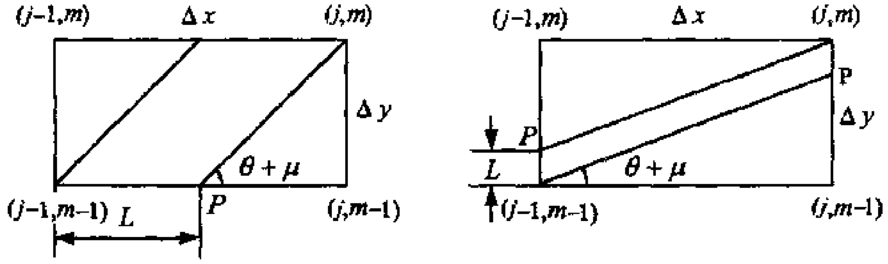


图 3.9

由于 $\phi_{j-1,m-1}$已知，则由 $M_{j-1,m-1}\rightarrow\mu_{j-1,m-1}\rightarrow\theta_{j-1,m-1}=\left(\arctan\frac{v}{u}\right)_{j-1,m-1}$ 均已知。其中 $\phi=(\rho,u,v,T,p)^T$，M 为马赫数，μ 为马赫角，θ 为流向角，u,v 为流向与法向分速。有两种可能：

(1) $\tan(\theta+\mu)_{j-1,m-1}>\frac{\Delta y}{\Delta x}$

过(j,m)作的特征线必然交于$(j-1,m-1)$点与$(j,m-1)$点的连线上。设交点为 P，则由线性插值得：$\phi_P=\phi_{j-1,m-1}+\frac{L}{\Delta x}(\phi_{j,m-1}-\phi_{j-1,m-1})$，于是 $\phi_{j,m}=\phi_P$，为此必须将 L 求出来。下面定 L：

定义 $\omega=\tan\left[\frac{\pi}{2}-(\mu+\theta)\right]$，则 $\omega_P=\tan\left[\frac{\pi}{2}-(\mu+\theta)_P\right]=\frac{\Delta x-L}{\Delta y}$。又 $\omega_P=\omega_{j-1,m-1}+\frac{L}{\Delta x}(\omega_{j,m-1}-\omega_{j-1,m-1})$，消去 ω_P 得：

$$L=\frac{\frac{\Delta x}{\Delta y}-\omega_{j-1,m-1}}{\frac{1}{\Delta x}(\omega_{j,m-1}-\omega_{j-1,m-1})+\frac{1}{\Delta y}} \tag{3.47a}$$

因为 $\omega_{j-1,m-1}$，$\omega_{j,m-1}$，Δx，Δy 已知，于是 $\phi_{j,m}$确定。

(2) 若 $\tan(\theta+\mu)_{j-1,m-1}<\frac{\Delta y}{\Delta x}$

则由(j,m)作出的特征线交于$(j-1,m-1)$点与$(j-1,m)$点的连线上。

用与上面类似的方法得：

$$L=\frac{\frac{\Delta y}{\Delta x}-\omega_{j-1,m-1}}{\frac{1}{\Delta y}(\omega_{j,m-1}-\omega_{j-1,m-1})+\frac{1}{\Delta y}} \tag{3.47b}$$

$$\phi_{j,m}=\phi_{j-1,m-1}+\frac{L}{\Delta y}(\phi_{j,m-1}-\phi_{j-1,m-1})$$

当然，可以用更精确些的方法处理上边界和外边界。

四、壁面附近网格加密

目前我们均采用二阶精度的差分格式解粘性流场。如 MacCormack 格式，Beam-Warming 格式及其他二阶精度格式如 TVD、NND 等，一阶精度格式满足不了要求，因此对流项要用中心差。前面在模型方程中已指出，二阶精度格式当 $Re(\Delta x)>2$ 后在物理量变化较大的地方必然要出现数值振荡。壁面附近就是这样的区域之一。解决的途径：

(1) 用一阶迎风格式，但物面附近正是物理粘性反应剧烈的地方，为使其不为数值耗散所淹没，不能用。二阶迎风格式就好些。

(2) 物面附近 $\left.\frac{\partial\phi}{\partial y}\right|_w=0$，各物理量无法向变化，这与实际情况不符，不能用。

(3) 保证物面附近 $Re\cdot\Delta l\leqslant 2$，至少应使 $Re\cdot\Delta l\sim O(1)$，当 $Re\approx 10^5\sim 10^6$；若 $\Delta y\sim\frac{1}{100}$（即 y 向分成 100 个网格），$Re\cdot\Delta y\approx 10^4\gg 10$。为保证 $Re\cdot\Delta y\approx 10$，必须 $\Delta y\approx\frac{1}{10^4}\sim\frac{1}{10^3}$。这时网格非常密，计算量和存储量都非常巨大。通常在粘性边界层，尤其是湍流边界层中，为使分辨率提高，$\Delta y\sim\frac{\delta}{10}$，$\delta$ 为边界层厚度。因此可用变网格间距的方法使壁面附近网格加密，远离壁面网格间距拉大，如图 3.10 所示。图中：

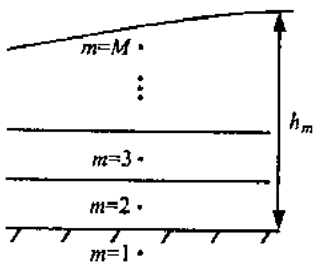


图 3.10

$$y_m=h_m\frac{e^{K[(m-3/2)/(M-1)]}-1}{e^K-1}。$$

h_m：壁面到上边界的距离；

m：网格编号；

K 值由 Re 与 y 向特征长度确定，保证当 $m=3/2$ 时 $y_m=0$。当 $m=5/2$ 时调整 K 可得要求的 $\Delta y_{min}\sim\frac{\sigma}{10}$。

可见从提高分辨能力与防止振荡上看均要求距壁面的第一个法向网格间距取得最小，然后不断延伸按指数规律放大。分析表明，当差分格式的精度提高时，壁面附近网格间距的限制可放宽。

3.7 抛物化 N. S. 方程的空间步进法

前面各节讲的方法主要用于时间相关法，即有时间项存在的问题，可充分利用初值问

题的求解技巧，统一地解出流场，但耗费机时，占用很多内存。它是全场迭代，目前尚不能满足复杂外形飞行器数值模拟的工程需要。

对大部分分离不太严重的问题，定常流场中总有一个主流方向存在（通常为 x 方向），各物理量沿主流方向耗散性的变化，与沿其他方向相比不占主要地位，那么可将沿主流方向粘性耗散变化忽略不计，因此，即使在亚声速时也可将方程化为抛物型。这就是抛物化 N. S. 方程的由来。抛物型方程的数学性质及相关问题已在第一章叙述。求解抛物型方程的定常解时不必加时间相关项，可使计算量和存储量大幅度减少。

3.7.1 抛物化 N. S. 方程的数学性质及定解条件的适定性

一、数学性质

将 N. S. 方程(1.1)去掉时间相关项，及去掉粘性耗散项中与主流方向导数有关的项得：

$$\frac{\partial F_i^*}{\partial x}+\frac{\partial G}{\partial y}=\frac{1}{\mathrm{Re}_*}\frac{\partial \bar{S}}{\partial y} \tag{1.16}$$

其中 $F_i^*=\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ (\rho E+p)u\end{bmatrix}$，$G=\begin{bmatrix}\rho v\\ \rho uv\\ \rho v^2+p\\ (\rho E+p)v\end{bmatrix}$，$\bar{S}=\begin{bmatrix}0\\ \mu\dfrac{\partial u}{\partial y}\\ \dfrac{4}{3}\mu\dfrac{\partial v}{\partial y}\\ \dfrac{\gamma k}{Pr}\dfrac{\partial E_1}{\partial y}+\mu u\dfrac{\partial u}{\partial y}+\dfrac{4}{3}\mu v\dfrac{\partial v}{\partial y}\end{bmatrix}$，

$E_1=\rho E=\dfrac{p}{\gamma-1}+\dfrac{\rho}{2}(u^2+v^2)$。

(1.16)的质量方程具有双曲型的性质，其他方程是抛物型的。但这只是数学上的分类。如果在计算中要利用抛物型方程的特点提初边值问题步进求解，还须满足一定条件。下面以(1.16)为例进行分析。

$$\frac{\partial F_i^*}{\partial U}\frac{\partial U}{\partial x}+\frac{\partial G}{\partial U}\frac{\partial U}{\partial y}=\frac{1}{\mathrm{Re}_*}\frac{\partial \bar{S}}{\partial\dfrac{\partial U}{\partial y}}\frac{\partial^2 U}{\partial y^2},U=(\rho,\rho u,\rho v,\rho E)^T$$

$$A\frac{\partial U}{\partial x}+B\frac{\partial U}{\partial y}=\frac{C}{\mathrm{Re}_*}\frac{\partial^2 U}{\partial y^2} \tag{3.49}$$

由(3.49)可见必须要求 Jacobian 矩阵 A 的特征值为正实数，矩阵 $A^{-1}C$ 的特征值有非负的实部，否则提初边值条件步进求解将是不适定的，这一适定性要求本来可以从数学上进行些解释，这里由应用角度，从物理概念上解说一下。双曲型方程$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}$反映了一个波的传播过程，如微弱扰动波的传播。若流场是光滑的，则这是一个可逆过程，因此，双曲型方程反映了可逆过程，而抛物型方程是热传导方程，是不可逆过程，因此在方程$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial y^2}$中，在逆主流方向的范围内提 Cauchy 问题是不适定的，热流只能由高温传导到低温。故上述抛物线方程中如 A 的一个特征值为负，将负值移到$\frac{\partial}{\partial x}$去，即是逆主流方向的传

播，于是出现不适定。计算也证明了这一点。$A^{-1}C$ 的特征值要求为正实部，主要是因为物理二阶耗散必须为正，这才使解达到稳定；否则如有负的二阶耗散，又不能用数值耗散将其淹没，必然导致解的不稳定而无法算下去。

二、抛物化 N.S. 方程的适定性分析

由(3.49)进行主变量的变换，即：

$$A\frac{\partial U}{\partial \phi}\frac{\partial \phi}{\partial x}+B\frac{\partial U}{\partial \phi}\frac{\partial \phi}{\partial y}=\frac{C}{\mathrm{Re}_*}\frac{\partial \frac{\partial U}{\partial y}}{\partial \frac{\partial \phi}{\partial y}}\frac{\partial^2 \phi}{\partial y^2}=\frac{C}{\mathrm{Re}_*}\frac{\partial U}{\partial \phi}\frac{\partial^2 \phi}{\partial y^2} \tag{3.50}$$

其中 $\phi=(\rho,u,v,p)^T$，$\frac{\partial U}{\partial \phi}=H=\begin{bmatrix}1 & 0 & 0 & 0\\ u & \rho & 0 & 0\\ v & 0 & \rho & 0\\ \frac{1}{2}(u^2+v^2) & \rho u & \rho v & \frac{1}{\gamma-1}\end{bmatrix}$，

$$H^{-1}=\begin{bmatrix}1 & 0 & 0 & 0\\ -\frac{u}{\rho} & \frac{1}{\rho} & 0 & 0\\ -\frac{v}{\rho} & 0 & \frac{1}{\rho} & 0\\ -\frac{\gamma-1}{2}(u^2+v^2) & (\gamma-1)u & -(\gamma-1) & \gamma-1\end{bmatrix}$$

以 H^{-1} 乘方程两侧得：

$$H^{-1}AH\frac{\partial \phi}{\partial x}+H^{-1}BH\frac{\partial \phi}{\partial y}=\frac{1}{\mathrm{Re}_*}H^{-1}CH\frac{\partial^2 \phi}{\partial y^2} \tag{3.51}$$

$$\bar{A}=H^{-1}AH=\begin{bmatrix}u & \rho & 0 & 0\\ 0 & u & 0 & \frac{1}{\rho}\\ 0 & 0 & u & 0\\ 0 & \gamma p & 0 & u\end{bmatrix}\text{，其特征值为 }\lambda_{\bar{A}}=\begin{cases}u+a\\ u-a\\ u\\ u\end{cases}$$

$\lambda_{\bar{A}}=\lambda_A$，即相似矩阵的特征值相等。

由此可知超声速时简化抛物型 N.S. 方程的特征值 $\lambda_A>0$，提初边值问题可能是适定的；而亚声速时即使抛物化了也因为 λ_A 中有一个值小于零而是不适定的，必须采用其他方法。下面分析矩阵 $A^{-1}C$ 的特征值：

$$A=\frac{\partial F}{\partial U}$$

$$=\begin{bmatrix}0 & 1 & 0 & 0\\ \frac{\gamma-3}{2}u^2+\frac{\gamma-1}{2}v^2 & -(\gamma-3)u & -(\gamma-1)v & \gamma-1\\ -uv & v & u & 0\\ u\left[\frac{(\gamma-2)(u^2+v^2)}{2}-\frac{a^2}{\gamma-1}\right] & \frac{(u^2+v^2)}{2}+(1-\gamma)u^2+\frac{a^2}{\gamma-1} & -(\gamma-1)uv & \gamma u\end{bmatrix} \tag{3.52}$$

求出 $A^{-1}C$ 后得其特征值为：

$$\lambda_{A^{-1}C}=\begin{cases}0\\ \dfrac{4}{3}\dfrac{\mu}{\rho u}\\ \dfrac{\mu}{2\rho(u^2-a^2)}\left[\left(u+\dfrac{\gamma u^2-a^2}{uN}\right)\pm\sqrt{\left(u+\dfrac{\gamma u^2-a^2}{uN}\right)^2-\dfrac{4\gamma(u^2-a^2)}{N}}\right]\end{cases}$$

其中$\dfrac{\gamma k}{P_r}=N>0$。令 $\beta=u+\dfrac{\gamma u^2-a^2}{uN}$，则 $A^{-1}C$ 的 4 个特征值为：

$$\lambda_1=0,\lambda_2=\frac{4}{3}\frac{\mu}{\rho u},\lambda_{3,4}=\frac{\mu}{2\rho(u^2-a^2)}\left[\beta\pm\sqrt{\beta^2-\frac{4\gamma(u^2-a^2)}{N}}\right] \tag{3.53}$$

这里 γ 是比热比，k 是传热系数，显然，当 $u>a$ 时，$\lambda_{A^{-1}C}\geqslant 0$；当 $u<a$ 时，λ_3 和 λ_4 中有一个取负值。

综上所述，抛物化 N. S. 方程具有抛物型的数学性质。超声速流动中提初边值问题进行空间步进是适定的；亚声速时如方法上不作改变，提初边值问题空间步进是不适定的，得不到稳定的解。

3.7.2 超声速抛物化方程的空间步进法

考虑$\dfrac{\partial F}{\partial x}+\dfrac{\partial G}{\partial y}=\dfrac{1}{\mathrm{Re}_\infty}\dfrac{\partial \bar{S}}{\partial y}$，用 MacCormack 显格式：

$$F_m^{\overline{j+1}}=F_m^j-\frac{\Delta x}{\Delta y}\left[(G_m^j-G_{m-1}^j)+\frac{1}{\mathrm{Re}_\infty}\left(\frac{U_{m+1}^j-U_m^j}{\Delta y}\right)-\frac{1}{\mathrm{Re}_\infty}\left(\frac{U_m^j-U_{m-1}^j}{\Delta y}\right)\right] \tag{3.54a}$$

$$F_m^{j+1}=\frac{1}{2}\left\{F_m^j+F_m^{\overline{j+1}}-\frac{\Delta x}{\Delta y}\cdot\left[\left(G_{m+1}^{\overline{j+1}}-G_m^{\overline{j+1}}\right)+\frac{1}{\mathrm{Re}_\infty}\left(\frac{U_{m+1}^{\overline{j+1}}-U_m^{\overline{j+1}}}{\Delta y}\right)-\frac{1}{\mathrm{Re}_\infty}\left(\frac{U_m^{\overline{j+1}}-U_{m-1}^{\overline{j+1}}}{\Delta y}\right)\right]\right\} \tag{3.54b}$$

由(3.54a)得：$F_m^{\overline{j+1}}=\begin{pmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ (\rho E+p)u\end{pmatrix}_m^{\overline{j+1}}=\begin{pmatrix}f_1\\ f_2\\ f_3\\ f_4\end{pmatrix}$，$f_1,f_2,f_3,f_4$ 已知，解得：

$$\begin{cases}v=\dfrac{f_3}{f_1}\\ u=\dfrac{\gamma}{\gamma+1}\left\{\dfrac{f_2}{f_1}\pm\sqrt{\left(\dfrac{f_2}{f_1}\right)^2-\dfrac{\gamma^2-1}{\gamma^2}\left[2\cdot\dfrac{f_4}{f_1}-\left(\dfrac{f_3}{f_1}\right)^2\right]}\right\}\\ p=f_2-f_1u\\ E=\dfrac{f_4-pu}{f_1}\\ \rho=\dfrac{f_1}{u}\end{cases} \tag{3.55}$$

于是 $G_m^{\overline{j+1}}$ 可按 $\begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2+p \\ (\rho E+p)v \end{pmatrix}_m^{\overline{j+1}}$ 求得，$U_m^{\overline{j+1}}$ 可按 $\begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix}_m^{\overline{j+1}}$ 求得。代入(3.54b)即可求解。

上述空间步进显格式的稳定性条件：

双曲－抛物方程 $A\dfrac{\partial U}{\partial x}+B\dfrac{\partial U}{\partial y}=C\dfrac{\partial^2 U}{\partial y^2}$ 可写成 $\dfrac{\partial U}{\partial x}+A^{-1}B\dfrac{\partial U}{\partial y}=A^{-1}C\dfrac{\partial^2 U}{\partial y^2}$，类似于(3.12)中 $G=0,\gamma_2=0$ 时的情况即式中 $\dfrac{\partial}{\partial x}\sim\dfrac{\partial}{\partial t},\dfrac{\partial}{\partial y}\sim\dfrac{\partial}{\partial x}$，其稳定性限制为：

$$\Delta t\leqslant\left(\frac{1}{\Delta t_{无粘}}+\frac{1}{\Delta t_{粘性}}\right)^{-1}=\frac{\Delta x^2}{2\bar{\gamma}+|\lambda_m|\Delta x} \tag{3.56}$$

上式即为(3.14)式的一维情形。当时已讲过，由于 $|\lambda_m|\Delta x\gg 2\bar{\gamma}$，因此可不计 $\bar{\gamma}$ 的影响。这里 $\bar{\gamma}$ 是运动粘性系数，注意与前面及后面 γ 所代表的比热比的区别。

将现在的问题化为 $\dfrac{\partial U}{\partial x}+A^{-1}B\dfrac{\partial \bar{u}}{\partial y}=\dfrac{A^{-1}C}{\mathrm{Re}_*}\dfrac{\partial^2 U}{\partial y^2}$，$\lambda_m=\lambda_{\max}(A^{-1}B)$。$A=\dfrac{\partial F}{\partial U}$ 见(3.52)式，$\lambda_{A^{-1}C}$ 已有(3.53)式，下面写出 B 矩阵及 $A^{-1}B$ 的特征值。

$$B=\begin{bmatrix} 0 & 0 & 1 & 0 \\ -uv & v & u & 0 \\ -v^2+\dfrac{\gamma-1}{2}(u^2+v^2) & -(\gamma-1)u & (3-\gamma)v & \gamma-1 \\ v[(\gamma-1)(u^2+v^2)-\gamma E] & -(\gamma-1)uv & \left[\gamma E-\dfrac{\gamma-1}{2}(u^2+3v^2)\right] & \gamma v \end{bmatrix} \tag{3.57}$$

由此求得：

$$\lambda_{A^{-1}B}=\begin{cases} \dfrac{v}{u} \\ \dfrac{v}{u} \\ \dfrac{(uv+a\sqrt{Q^2-a^2})}{u^2-a^2} \\ \dfrac{(uv-a\sqrt{Q^2-a^2})}{u^2-a^2} \end{cases} \tag{3.58}$$

$$Q^2=u^2+v^2$$

在超声速情况下 $\lambda_{A^{-1}B}$ 均为实数。令(3.56)式中 $\Delta t\sim\Delta x,\Delta x\sim\Delta y$，则：

$$\Delta x\leqslant\frac{\Delta y}{|\lambda_{A^{-1}B}|_{\max}}=\frac{(u^2-a^2)\Delta y}{uv+a\sqrt{Q^2-a^2}}。$$

通常粘性项影响不大。如要考虑则：

$$\Delta x\leqslant\frac{\Delta y^2}{2|\lambda_{A^{-1}C}|_{\max}+\dfrac{uv+a\sqrt{Q^2-a^2}}{(u^2-a^2)}\Delta y} \tag{3.59}$$

由此可见，在 $M\approx 1$ 附近，$u^2\approx a^2$ 要求 Δx 十分小，要设法避免。

3.7.3 亚声速流场抛物化 N.S. 方程的空间步进法

亚声速抛物化N.S.方程简化后空间步进的不适定性可先从物理上来了解一下，以定常的 x 向动量方程为例：$\frac{\partial u}{\partial x}+\frac{v}{u}\frac{\partial u}{\partial y}=-\frac{1}{\rho u}\frac{\partial p}{\partial x}+\frac{\mu}{\rho u}\frac{\partial^2 u}{\partial y^2}$。所谓空间步进，就是 $u_m^{j+1}=u_m^j+[\cdots]^j$，因此 $(j+1)\Delta x$ 处的物理量 u 及其他量均受其上游的各物理量的影响，下游各物理量不影响它，最多只通过上、下边界量的调节进行间接影响。

超声速流动时，流场各参数所受到的影响可用前马赫锥来表示，影响是单侧的见图 3.11，因此空间步进是可行的，适定的。亚声速流动时后面的物理量对前面也有影响，若仍用上述空间步进而不计后面的影响，不作修正显然是不合适的，从而解不适定。

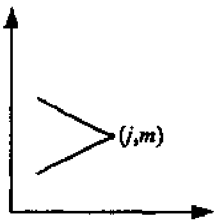


图 3.11

从动量方程可见"$\frac{\partial p}{\partial x}$"是产生不适定性的根本原因。正确处理好压力项是求解 P.N.S. 方程的关键，而且 p 又受到 u、v、ρ、T 的综合影响。

一、设$\frac{\partial p}{\partial x}=0$

此时求 A_p 矩阵时可设 p 为常数。

$$A_p=\frac{\partial F}{\partial U}=\frac{\partial[\rho u,p+\rho u^2,\rho uv,(\rho E+p)u]^T}{\partial[\rho,\rho u,\rho v,\rho E]^T}$$

$$=\begin{bmatrix} 0 & 1 & 0 & 0 \\ -u^2 & 2u & 0 & 0 \\ -uv & v & u & 0 \\ -(\rho E+p)\frac{u}{\rho} & \frac{1}{\rho}(\rho E+p) & 0 & u \end{bmatrix} \tag{3.60}$$

求得其特征值为 $\lambda_{A_p}=(u,u,u,u)^T$。只要具有主流方向的速度，则特征值恒正，满足适定性的要求。对粘性项 $A_p^{-1}C$，

$$\lambda_{C_p}=\frac{\mu}{\rho u}\left(0,1,\frac{4}{3},\frac{\gamma k}{\mu P_r}\right)^T \tag{3.61}$$

均为非负的实数，故粘性耗散项在步进时起正耗散的稳定作用，从而可大胆地在亚声速区内用空间步进法求解。从物理上说，由于$\frac{\partial p}{\partial x}=0$，保证下游的影响不可能通过 p 而上传，而其他物理量均不含有前传效应，因此按纯抛物型的方程一层层往下算，得到稳定的结果。若实际流动中$\frac{\partial p}{\partial x}\neq 0$，则这种结果意义不大。

二、部分压力修正法

从前面关于亚声速流 P.N.S. 方程空间步进的分析得知，不适定的问题主要是由 $u<a$ 引起的，而通过对 p 的限制可改变其系数矩阵的特征值，但不必限制 p 为常数，只需确保小于 0 的那个特征值改变为正数。

令
$$\left(\frac{\partial p}{\partial x}\right)_m^j=\omega\left(\frac{\partial p}{\partial x}\right)_m^j+(1-\omega)\left(\frac{\partial p}{\partial x}\right)_m^{j-1} \tag{3.62}$$

其中$(j-1)$层是上一次算出来的已知值，ω 是小于 1 的可调参数。于是：

$$\frac{\partial F}{\partial U}=A_\omega$$

$$=\begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{\omega}{2}(\gamma-1)(u^2+v^2)-u^2 & 2u-\omega u(\gamma-1) & -(\gamma-1)\omega v & (\gamma-1)\omega \\ -uv & v & u & 0 \\ u[(\gamma-1)(u^2+v^2)-\gamma E] & \gamma E-\frac{3}{2}(\gamma-1)u^2-\frac{\gamma-1}{2}v^2 & (\gamma-1)uv & \gamma u \end{bmatrix} \tag{3.63}$$

由此求得：

$$\lambda_{A\omega}=\begin{cases} u-a\sqrt{\dfrac{\omega}{\gamma-(\gamma-1)\omega}} \\ u+a\sqrt{\dfrac{\omega}{\gamma-(\gamma-1)\omega}} \\ u \\ u \end{cases}$$

为保证 $u<a$ 时 $u-a\sqrt{\frac{\omega}{\gamma-(\gamma-1)\omega}}\geqslant 0$ 必须：

$$\omega\leqslant\frac{\gamma M_x^2}{1+(\gamma-1)M_x^2},M_x=\frac{u}{a} \tag{3.64}$$

此法的缺点是，人为地用上一步结果的某一部分来代替当前步求$\frac{\partial p}{\partial x}$的值，而且其所影响的程度由当地 M 数确定，因此有一定近似性，对局部亚声速流区有效。此外还有修正速度或修正密度的方法，目的均是人为地近似给出分布，使 A 的特征值在亚声速时不小于零。

3.7.4 Spalding-Patankar 方法空间步进解亚声速流场

一、Spalding-Patanker 法计算要点

用空间步进法解轴对称或二维亚声速粘性流场是一个吸引人的课题。从原理上说前面已大致解释了，可以用各种近似方法数值求解。Spalding-Patanker 解热传导与流动问题上的一系列方法可作为借鉴。现在具体介绍一些规则和要求，其特点是非守恒的差分方程求解。

1. 数值离散规则

以一维常微分方程为例：

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(k\frac{\mathrm{d}T}{\mathrm{d}x}\right)+S=0$$

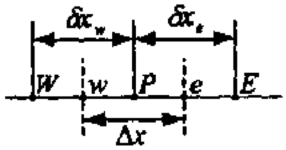


图 3.12

图 3.12 上虚线之间表示控制长度；在 Δx 控制长度上积分此方程得：

$$\left(k\frac{\mathrm{d}T}{\mathrm{d}x}\right)_e-\left(k\frac{\mathrm{d}T}{\mathrm{d}x}\right)_w+\int_w^e S\mathrm{d}x=0$$

$$\frac{k_e(T_E-T_P)}{\delta x_e}-\frac{k_w(T_P-T_W)}{\delta x_w}+\overline{S}\Delta x=0$$

$\overline{S}$ 是 S 在 Δx 上的平均值。所以可整理为：

$$a_P T_P=a_E T_E+a_W T_W+b=\sum a_{nb}T_{nb}+b$$

式中 $a_E=\dfrac{k_e}{\delta x_e}$，$a_W=\dfrac{k_w}{\delta x_w}$，$a_P=a_E+a_W$，$b=\overline{S}\cdot\Delta x$，下标 P 是指中心点上的物理量，E，W 是指其东西两侧的节点上的物理量，故用 $a_{nb}T_{nb}$ 表示之。这样的形式对二、三维同样适用，只是 $a_{nb}T_{nb}$ 表示上下左右各节点的值而已。它们常常是用非守恒方程的非守恒离散，离散时必须注意如下四个基本规则：

(1) 交界面上物理量的流通要相容

如图 3.12 中的 w 面既属于 P 又属于 W，在 w 处流入与流出的热流必须用相同的表达式，即保持守恒性。这种约定常常被无意识地违背。例如：

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(k\frac{\mathrm{d}T}{\mathrm{d}x}\right)=\frac{k_P(T_E-T_P)}{\delta x_e}-\frac{k_P(T_P-T_W)}{\delta x_w}$$

此时从以 P 为中心的控制体流出 e 面的热流量是 $\dfrac{k_P(T_E-T_P)}{\delta x_e}$，若下一个以 E 为中心的控制体 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(k\dfrac{\mathrm{d}T}{\mathrm{d}x}\right)=\dfrac{k_E(T_F-T_E)}{\delta x_f}-\dfrac{k_E(T_E-T_P)}{\delta x_e}$ 流入 e 面的是 $\dfrac{k_E(T_E-T_P)}{\delta x_e}$，$k_E\neq k_P$ 二者就不相容了。另外，如用二次函数分布，T 也会出现类似问题。尤其是对流项 $v\dfrac{\partial u}{\partial y}$ 的"v"的取法很有讲究，原则：交界面上的流量属于交界面本身，因此，写非守恒形式时，"v"应取交界面上的值。

(2) 要求正的系数

即所有系数 a_{nb} 与 a_P 必须总是正的，尤其是考虑对流项后这个准则十分重要。其他条件不变，耗散与对流影响的结果使一个节点物理量的增加必然使邻近节点物理量相应增加而不是减少，即 a_{nb} 与 a_P 必须同符号。

(3) 源项与线化规定

源项 $\overline{S}=S_c+S_P T_P$，这是源项与变量有关进行线化的一种规则，通常要求 $S_P\leqslant 0$。因为 $a_P T_P=\sum a_{nb}T_{nb}+(S_C+S_P T_P)\Delta x$，所以 $(a_P-S_P\Delta x)T_P=\sum a_{nb}T_{nb}+S_C\Delta x=\bar{a}_P T_P$。

当 S_P 较大时 T_{nb} 的增加也可能使 T_P 下降，因 $\bar{a}_P$ 可能为负值；而 $S_P<0$ 保证 $\bar{a}_P>0$，常常可使不稳定及物理上的不真实解不会出现，所以 $S_P<0$ 是根本的。

(3) 邻近系数 a_{nb} 之和必等于 a_P。如上面的微分方程中 T 是方程的解，$(T+C)$ 也必是方程的解。离散化方程中要实现这一点，必要求邻近系数 a_{nb} 之和等于 a_P。但当源项也与 T 有关时不一定遵守此原则。

当我们写好离散化方程时，引用上述规则判断一下是十分必要的。

2. 二维流场求解的网格处理

有两种可借用的网格处理法：A 法，控制容积面位于两节点的中间(图 3.13)；B 法，节点在控制容积的中心(图 3.14)。

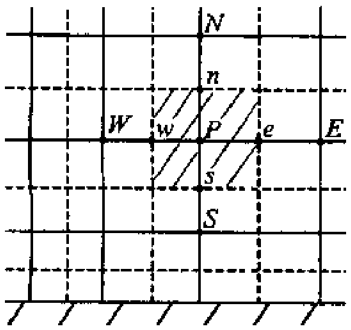


图 3.13 非等间距网格，P 点不在有虚线组成的控制面的中心

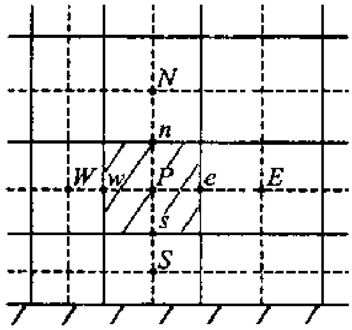


图 3.14 非等间距网络，P 点位于由实线组成的控制面中心，但控制面不在 P、E 两节点中心

讨论：均匀网格二者是一致的；对变间距网格来说，B 法优点较多，因为控制面是建立离散的基本单元，以其中心点处的 P 代表该面内平均物理量显然是方便的、正确的；考虑边界条件时，A 法的控制容积要另外考虑（半容积），而 B 法都用正常的控制容积考虑，并且将边界值布置在靠近边界的容积的节点上可采用平均的办法灵活处理边界条件。

3. 交错网格

问题的提出：在动量方程中存在 $-\frac{\partial p}{\partial x}$ 项，这是影响计算的重要项，下面用控制容积的方法：

$$-\frac{\partial p}{\partial x}=-\frac{p_e-p_w}{\delta x}=\frac{\frac{p_W+p_P}{2}-\frac{p_P+p_E}{2}}{\delta_x}=\frac{p_W-p_E}{2\Delta x}$$

这意味着动量方程中物理量的变化将受两个相间隔节点之间压力差的影响。

当 p 是图 3.15 所示锯齿形分布时，在动量方程的离散方程中，意味着没有压力场的作用，这个情况将会一直到计算结束，显然是不真实的，尤其当二维流场出现锯齿棋盘式分布时，情况更是如此。如果实际分布是连续的，上述说明意味着加一个棋盘锯齿分布的压力场不影响计算结果，这显然是谬误的。这是离散的方式所引起的问题，在质量方程中速度场的锯齿形分布也有类似不合理情况。

解决这个困难的办法是不同变量布置在不同的位置上，即所谓“交错”网格。规定各速度分量都布置在控制容积表面的中心处，法向分速布置在其垂直流过的表面中间，切向分速亦相应布置，而 ρ、T、p 等均布置在节点处。其优点之一，即使有节点上的锯齿速度场，但由于质量方程可用邻近两点处流出流进的质量流率表示，可保证不会使锯齿形结果作为均匀流量来处理；同时相邻两点锯齿形的压力场再也不会使速度均匀，因为有交界面处的速度分量来平衡此压力差，于是随计算的进行锯齿形状态将被打破。当然，用了交错网格，计算上会带来插值等的麻烦，但这是值得的。

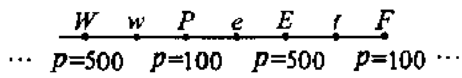


图 3.15

二、控制方程及其离散差分格式

输运型方程：$\frac{\partial\rho\phi}{\partial t}+\frac{\partial}{\partial x_j}(\rho u_j\phi)=\frac{\partial}{\partial x_j}\left(\Gamma\frac{\partial\phi}{\partial x_j}\right)+S,\Gamma=\left(0,\mu,\frac{4}{3}\mu,\frac{\gamma k}{P_r}\right)^T,\phi=(1,u,v,E)^T$；$u_j=u,j=1;u_j=v,j=2;x_j=x,j=1;x_j=y,j=2$；将$\frac{\partial p}{\partial x},\frac{\partial p}{\partial y},\mu u\frac{\partial u}{\partial y}$等项作为源项加于 S 中。这种对流耗散方程能代表一般形式的 N. S. 方程。下面取一维定常模型方程 $\frac{\mathrm{d}}{\mathrm{d}x}(\rho u\phi)=\frac{\mathrm{d}}{\mathrm{d}x}\left(\Gamma\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)$进行分析。

1. 中心差分

如图 3.16，$(\rho u\phi)_e-(\rho u\phi)_w=\Gamma_e\left(\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)_e-\Gamma_w\left(\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)_w$.

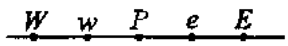


图 3.16

一维流连续方程 ρu=常数。又 $\phi_e=\frac{\phi_E+\phi_P}{2},\phi_w=\frac{\phi_P+\phi_W}{2}$，所以：

$$\frac{(\rho u)_e}{2}(\phi_P+\phi_E)-\frac{(\rho u)_w}{2}(\phi_P+\phi_W)=\frac{\Gamma_e}{\delta x_e}(\phi_E-\phi_P)-\frac{\Gamma_w}{\delta x_w}(\phi_P-\phi_W)$$

$$a_P\phi_P=a_E\phi_E+a_W\phi_W$$

$$a_E=\frac{\Gamma_e}{\delta x_e}-\frac{(\rho u)_e}{2}=D_e-\frac{F_e}{2},\quad a_W=\frac{\Gamma_w}{\delta x_w}-\frac{(\rho u)_w}{2}=D_w+\frac{F_w}{2}$$

$$a_P=D_e-\frac{F_e}{2}+D_w+\frac{F_w}{2}$$

其中 $D_e=\frac{\Gamma_e}{\delta x_e},F_e=(\rho u)_e$，由前面四规则要求 $a_P=a_E+a_W$，这显然均成立；要求 a_P,a_E,a_W 同号，$a_E=D_e-\frac{F_e}{2}=D_e\left[1-\frac{(\rho u)_e\delta x_e}{2\Gamma_e}\right]=D_e\left[1-\frac{P_C}{2}\right]$，$P_C$ 称为 Peclet 数（其实 Γ_e 相当于粘性系数 μ_e，P_C 相当于网格雷诺数）；当 $P_C>2$ 后，a_E 可能小于 0，违背原则，故大 Re 时采用这种格式是不太合适的。事实上，认为 $\phi_e=\frac{\phi_E+\phi_P}{2}$，即 $\phi\sim x$ 是线性分布的结果，也不太合适。

2. 迎风格式

对流项的差分法：当 $(\rho u)_e<0$ 时，$\phi_e=\phi_E$；当 $(\rho u)_e>0$ 时，$\phi_e=\phi_P$；于是：$[\phi_P\{F_e,0\}-\phi_E\{-F_e,0\}]-[\phi_W\{F_w,0\}-\phi_P\{-F_w,0\}]=\frac{\Gamma_e}{\delta x_e}(\phi_E-\phi_P)-\frac{\Gamma_w}{\delta x_w}(\phi_P-\phi_W)$。其中$\{a,b\}=\mathrm{Max}\{a,b\}$。所以：

$$a_P\phi_P=a_E\phi_E+a_W\phi_W$$

$$a_P=D_e+\{F_e,0\}+D_w+\{-F_w,0\},$$

$$a_E=D_e+\{-F_e,0\},a_W=D_w+\{F_w,0\}$$

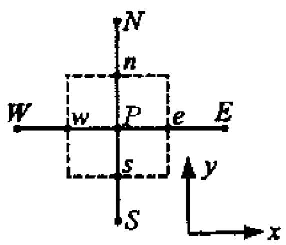


图 3.17

显然因 $F_e=F_w$，$a_P=a_E+a_W$；同时，$a_E>0,a_W>0$，故满足了基本准则，但能否采用要进一步分析；大 Re 时人工粘性过大，不合要求。

另外还可采用混合格式及半解析格式等进行离散，这里就不详述了。

三、二维控制方程的离散

$$\frac{\partial\rho\phi}{\partial t}+\frac{\partial J_x}{\partial x}+\frac{\partial J_y}{\partial y}=S$$

其中 $J_x=\rho u\phi-\Gamma\frac{\partial\phi}{\partial x}$，$J_y=\rho v\phi-\Gamma\frac{\partial\phi}{\partial y}$。

由图 3.17 积分得：

$$\frac{\rho_P\phi_P-\rho_P^0\phi_P^0}{\Delta t}\Delta x\cdot\Delta y+\Delta y(J_e-J_w)+\Delta x(J_n-J_s)=(S_C+S_P\phi_P)\Delta x\cdot\Delta y \tag{3.65}$$

“0”表示 $n\Delta t$ 时间的值，用全隐格式，不带“0”的均指“$n+1$”时的值。

同时满足质量方程：

$$\frac{\rho_P-\rho_P^0}{\Delta t}\Delta x\cdot\Delta y+F_e-F_w+F_n-F_s \tag{3.66}$$

$$F_e=(\rho u)_e\Delta y, F_w=(\rho u)_w\Delta y, F_n=(\rho u)_n\Delta x, F_s=(\rho u)_s\Delta x$$

(3.66)$\times\phi_P-$(3.65)得：

$$\begin{aligned}&\frac{(\phi_P-\phi_P^0)\rho_P^0}{\Delta t}\Delta x\cdot\Delta y+(J_e-F_e\phi_P)\\&-(J_w-F_w\phi_P)+(J_n-F_n\phi_P)-(J_s-F_s\phi_P)\\&=(S_C+S_P\phi_P)\Delta x\cdot\Delta y\end{aligned}$$

其中

$$J_e-F_e\phi_P=\left(\rho u\phi-\Gamma\frac{\partial\phi}{\partial x}\right)_e-(\rho u)_e\phi_P=a_E(\phi_P-\phi_E),$$

$$a_E=D_e\left\{0,1-\frac{|P_{\alpha}|}{2}\right\}+\{-F_e,0\}。$$

同理

$$J_w-F_w\phi_P=a_W(\phi_W-\phi_P), a_W=D_w\left\{0,1-\frac{|P_{cw}|}{2}\right\}+\{F_w,0\},\cdots\cdots$$

最后，差分方程为：

$$a_P\phi_P=a_E\phi_E+a_W\phi_W+a_N\phi_N+a_S\phi_S+b=\sum a_{nb}\phi_{nb}+b$$

$$a_P=a_E+a_W+a_N+a_S+a_P^0+S_P\Delta x\cdot\Delta y$$

$$a_P^0=\frac{\rho_P^0\Delta x\cdot\Delta y}{\Delta t}, b=S_C\Delta x\cdot\Delta y+a_P^0\phi_P^0,$$

$$D_e=\frac{\Gamma_e\Delta y}{\delta x_e}, D_w=\frac{\Gamma_w\Delta y}{\delta x_w}, D_n=\frac{\Gamma_n\Delta x}{\delta x_n}, D_s=\frac{\Gamma_s\Delta x}{\delta x_s}, P_{ce}=\frac{(\rho u)_e\delta x_e}{\Gamma_e}, P_{cw}=\frac{(\rho u)_w\delta x_w}{\Gamma_w}。$$

注意将时间相关项与其他隐式项归并于 a_P 中一起处理。

四、SIMPLE 法求解抛物化 N.S. 方程组

1. 压力修正方程

按已讲过的办法将动量方程等分开求解，但源项中压力项要专门处理。令 $u=u^*+u'$，$v=v^*+v'$，$p=p^*+p'$，其中“$*$”为平均值，“$'$”为扰动值，则 x 向动量方程为：

$$a_e u_e^*=\sum a_{nb}u_{nb}^*+b+(p_P^*-p_E^*)A_e \tag{3.67}$$

注意：u 布置在 e 处(交错网格，控制容积为 A_e)；p_E^* 是猜测值，u_e 是要求解的预测值；其他均是已知值。

同理，y 向动量方程为：

$$a_n v_n^* = \sum a_{nb} v_{nb}^* + b + (p_P^* - p_N^*) A_n$$

原式 $a_e u_e = \sum a_{nb} u_{nb} + b + (p_P - p_E) A_e$，与(3.67)相减得：$a_e u_e' = (p_p' - p_E') A_e$，于是：

$$u_e = u_e^* + d_e(p_P' - p_E'), \quad d_e = A_e / a_e \tag{3.68a}$$

同理

$$v_n = v_n^* + d_n(p_P' - p_N'), \quad d_n = A_n / a_n \tag{3.68b}$$

因质量守恒有：

$$\frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} = 0$$

$$[(\rho u)_e - (\rho u)_w]\Delta y + [(\rho v)_n - (\rho v)_s]\Delta x = 0$$

将(3.68)代入，得：

$$a_P p_P' = a_E p_E' + a_W p_W' + a_N p_N' + a_S p_S' + b \tag{3.69}$$

其中：$a_E = \rho_e d_e \Delta y$，$a_W = \rho_w d_w \Delta y$，$a_N = \rho_n d_n \Delta x$，$a_S = \rho_s d_s \Delta x$，$a_P = a_E + a_W + a_N + a_S$，$b = [(\rho u^*)_w - (\rho u^*)_e]\Delta y + [(\rho v^*)_s - (\rho v^*)_n]\Delta x$。$\rho$ 值分布在节点上，可用线性插值法给出。这就是压力校正方程。

2. SIMPLE 压力连接方程的半隐式求解

SIMPLE 是 Semi-Implicit Method for Pressure Linked Equations 的简写，求解过程如下：

(1) 试猜压力场 p^*；

(2) 动量方程求预测值 u^*, v^*；

(3) 求解压力校正方程得 p'；

(4) 求 $p = p^* + p'$；

(5) 求 u_e, v_n；见(3.68)式。

(6) 求 T 等其他参数；

(7) 将 p 作为一个新的 p^* 重复(2)～(6)，直到解收敛。

3. 讨论

(1) 为了使 p' 的校正方程直接导出，将速度修正方程中 $\sum a_{nb} u_{nb}'$，$\sum a_{nb} v_{nb}'$ 均略去，说明校正压力对邻点速度的影响不考虑，因此隐式是不完全的，故称为半隐式。这样的略去是允许的，因为只要解收敛，最后迭代得 $p' = 0, b = 0$，不会改变最后结果；

(2) 实际计算中通常由 $b \approx 10^{-5} \sim 10^{-6}$ 作为收敛的判断；

(3) 解压力校正方程与动量方程时均要用欠松弛迭代。这个问题在椭圆型方程的解法中有介绍。

第四章　应用计算流体力学专题

20 世纪 80 年代以来，随着计算流体力学在航空、航天及其他工程领域复杂流场计算中的广泛应用，应用计算流体力学得到了蓬勃的发展，计算方法及网格生成技术等均有了巨大的改进。本章拟以专题的形式对其中的部分内容加以介绍，包括二阶精度的总变差递减(TVD)格式的机理；高阶基本无振荡(ENO)格式的分析研究；定常流数值计算时加快收敛的措施，如 LU-SSOR、LU-SGS 方法，人工压缩性法及预处理法(Precondition)，非结构网格及分区结构网格生成等。

4.1　二阶 TVD 格式的分析

第二章中曾介绍过总变差递减(TVD)格式，主要说明了 NND 格式和构成 TVD 格式的条件。在上述内容的基础上，本章将对 TVD 格式的机理作进一步的分析。

4.1.1　由一阶迎风格式构成二阶 TVD 格式

对于微分方程：

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0 \tag{1.58}$$

其中 $a>0$。上述方程的一阶迎风格式为：

$$u_j^{n+1}=u_j^n-\sigma(u_j^n-u_{j-1}^n) \tag{1.63a}$$

前已证明，这是符合 TVD 条件的，但过激波抹平太强，反映正耗散过多。

(1.58)的 Lax-Wendroff 格式为：

$$u_j^{n+1}=u_j^n-\frac{\sigma}{2}(u_{j+1}^n-u_{j-1}^n)+\frac{\sigma^2}{2}(u_{j+1}^n-2u_j^n+u_{j-1}^n) \tag{2.29}$$

可改写为：

$$u_j^{n+1}=u_j^n(1-\sigma^2)-(\frac{\sigma}{2}-\frac{\sigma^2}{2})u_{j+1}^n+(\frac{\sigma}{2}+\frac{\sigma^2}{2})u_{j-1}^n$$

前已证明，这不满足 TVD 条件。现将上式改写成下面的形式：

$$u_j^{n+1}=u_j^n-\sigma(u_j^n-u_{j-1}^n)-\frac{\sigma}{2}(1-\sigma)(u_{j+1}^n-2u_j^n+u_{j-1}^n) \tag{4.1}$$

此式右侧前两项即构成一阶迎风格式。右侧第三项用 Taylor 级数展开可得为：

$$\frac{\sigma}{2}(1-\sigma)\frac{\partial^2 u}{\partial x^2}\bigg|_j^n\Delta x^2+\cdots$$

可见这是由一阶迎风格式减去一个$\frac{\sigma}{2}(1-\sigma)\frac{\partial^2 u}{\partial x^2}\Big|_j^n\Delta x^2+\cdots$(其首项为二次耗散的截断误差)构成的，即一阶迎风二阶耗散过大，而 Lax-Wendroff 格式是减去的耗散过多，致使格

式精度虽提高但 TVD 性能丧失，能否在 Lax-Wendroff 格式(4.1)的基础上少减去一些二阶耗散，使得既保证 TVD，又保持二阶精度，事实上是可以的。

一、u 的变化率

要求 u_j^n 的变化是单调的，如何通过限制负耗散来达到，人们发现可用 u_j^n 的变化率来进行限制。具体作法，由前面的 Lax-Wendroff 格式(4.1)可得：

$$u_j^{n+1} = u_j^n - \sigma\Delta u_{j-1}^n - q(\Delta u_j^n - \Delta u_{j-1}^n) \tag{4.2}$$

其中 $q=\frac{\sigma}{2}(1-\sigma)$，$\Delta u_j=u_{j+1}-u_j$，$\Delta u_{j-1}=u_j-u_{j-1}$。

用变化率限制，设 $\eta_{j+1/2}=\frac{\Delta u_{j-1}}{\Delta u_j}$，代入

$$u_j^{n+1} = u_j^n - \left[\sigma + q\left(\frac{1}{\eta_{j+1/2}} - 1\right)\right]\Delta u_{j-1} \tag{4.3}$$

令 $s=\sigma+q\left(\frac{1}{\eta_{j+1/2}}-1\right)$，且有 $0\leqslant s\leqslant 1$，则可满足 TVD 性质，此时：

$$u_j^{n+1} = u_j^n - s(u_j^n - u_{j-1}^n) = (1-s)u_j^n + su_{j-1}^n \tag{4.4}$$

由(2.138)知：

$$c_{j+\frac{1}{2}} = 0, 1 - c_{j+\frac{1}{2}} - c_{j-\frac{1}{2}} = 1 - s \geqslant 0, c_{j-\frac{1}{2}} = s \geqslant 0$$

因此满足 TVD 条件。

过去对 s 无限制，使得 Lax-Wendroff 格式没有 TVD 性质，即对一阶迎风格式反耗散过多。通过上述分析有：

$$0\leqslant s\leqslant 1, \text{即 } 0\leqslant \sigma + \frac{\sigma}{2}(1-\sigma)\left(\frac{1}{\eta} - 1\right) \leqslant 1 \tag{4.5}$$

所以

$$1 + \frac{1-\sigma}{2}\left(\frac{1}{\eta} - 1\right) \geqslant 0, \frac{1-\sigma}{2}\left(\frac{1}{\eta} - 1\right) \geqslant -1, \frac{1}{\eta} - 1 \geqslant \frac{-2}{1-\sigma}$$

得：

$$\eta \geqslant \frac{\sigma - 1}{1 + \sigma} \tag{4.6}$$

又 $1 + \frac{1-\sigma}{2}\left(\frac{1}{\eta} - 1\right) \leqslant \frac{1}{\sigma}$，得：

$$\eta \geqslant \frac{\sigma}{2+\sigma} \tag{4.7}$$

最终得 $\eta_{j+\frac{1}{2}} \geqslant \frac{\sigma}{2+\sigma}$。

二、限制器

现在分析上述对 s 的限制可通过对(4.3)引入速度变量限制器 $\psi(\eta)$ 来达到：

$$\begin{aligned} u_j^{n+1} &= u_j^n - \sigma\Delta u_{j-1}^n - q(\psi_{j+\frac{1}{2}}\Delta u_j^n - \eta_{j-\frac{1}{2}}\Delta u_{j-1}^n) \\ &= u_j^n - \left[\sigma + q\left(\frac{\psi_{j+\frac{1}{2}}}{\eta_{j+\frac{1}{2}}} - \psi_{j-\frac{1}{2}}\right)\right]\Delta u_{j-1}^n \end{aligned} \tag{4.8}$$

令：

$$Q=\left[\sigma+q\left(\frac{\psi_{j+\frac{1}{2}}}{\eta_{j+\frac{1}{2}}}-\psi_{j-\frac{1}{2}}\right)\right] \tag{4.9}$$

如(4.3)要求 $0\leqslant Q\leqslant 1$，才能满足 TVD 条件。

令 $Q_1=\dfrac{\psi_{j+\frac{1}{2}}}{\eta_{j+\frac{1}{2}}}-\psi_{j-\frac{1}{2}}$，则由(4.9)知：

$$0\leqslant(\sigma+qQ_1)\leqslant 1 \tag{4.10}$$

从而

$$\frac{-2}{1-\sigma}\leqslant Q_1\leqslant\frac{2}{\sigma}，且\ \sigma\leqslant 1 \tag{4.11}$$

控制器 ψ 与 Q 之关系是

$$Q_1=\frac{\psi_{j+\frac{1}{2}}}{\eta_{j+\frac{1}{2}}}-\psi_{j-\frac{1}{2}} \tag{4.12}$$

因为 $\eta\geqslant 0$，且由(4.7)$\eta\geqslant\dfrac{\sigma}{2+\sigma}$，如限制 $Q\geqslant 0$，即 $\eta\geqslant 0,\psi(\eta)\geqslant 0,\dfrac{\psi}{\eta}-\psi\leqslant\dfrac{2}{\sigma}$，于是 TVD 要求变成：

$$0\leqslant\frac{\psi(\eta)}{\eta}\leqslant\frac{2}{\sigma}\ 及\ 0\leqslant\psi(\eta)\leqslant\frac{2}{1-\sigma} \tag{4.13}$$

现分析取不同限制器 $\psi(\eta)$ 的情况：

(1) $\psi(\eta)=1$

$$\begin{aligned}u_j^{n+1}&=u_j^n-\sigma\Delta u_{j-1}-\frac{\sigma(1-\sigma)}{2}(\Delta u_j-\Delta u_{j-1})\\&=u_j^n-\sigma\Delta u_{j-1}-\frac{\sigma(1-\sigma)}{2}(u_{j+1}^n-2u_j^n+u_{j-1}^n)\end{aligned}$$

即 Lax-Wendroff 格式，具有二阶精度，但是由$\dfrac{\psi(\eta)}{\eta}\leqslant\dfrac{2}{\sigma}$，$\eta=\dfrac{\sigma}{2+\sigma}$知$\dfrac{2+\sigma}{\sigma}\leqslant\dfrac{2}{\sigma}$不成立，不满足(4.13)式，因此 $\psi(\eta)=1$ 时不是 TVD 格式。

(2) $\psi(\eta)=\eta$

$$\begin{aligned}u_j^{n+1}&=u_j^n-\left[\sigma+q\left(\frac{\psi_{j+1/2}}{\eta_{j+1/2}}-\psi_{j-1/2}\right)\right]\Delta u_{j-1}\\&=u_j^n-[\sigma+q(1-\eta_{j-1/2})]\Delta u_{j-1}\end{aligned}$$

其中

$$\Delta u_{j-1}=u_j-u_{j-1},\eta_{j-1/2}=\frac{\Delta u_{j-2}}{\Delta u_{j-1}}$$

$$\begin{aligned}u_j^{n+1}&=u_j^n-\sigma(u_j^n-u_{j-1}^n)-\frac{\sigma}{2}(1-\sigma)(u_j^n-u_{j-1}^n)+\frac{\sigma}{2}(1-\sigma)(u_{j-1}^n-u_{j-2}^n)\\&=u_j^n-\sigma(u_j^n-u_{j-1}^n)-\frac{\sigma}{2}(1-\sigma)(u_j^n-2u_{j-1}^n+u_{j-2}^n)\end{aligned}$$

上式是二阶迎风格式，具有二阶精度，且有 $0\leqslant\dfrac{\psi(\eta)}{\eta}=1\leqslant\dfrac{2}{\sigma}$。此时，$\sigma\leqslant 2$ 是其稳定性

条件。但 $\psi=\eta\leqslant\frac{2}{1-\sigma}$，故当 $1<\sigma\leqslant2$ 时，η 为负值，与 $\eta\geqslant0$ 的要求有矛盾，故二阶迎风格式在 $1<\sigma\leqslant2$ 时不是 TVD 格式。

从上面讨论可见，当 ψ 与 η 的关系为 $1\geqslant\psi\geqslant\eta$ 时，(4.8) 式 $u_j^{n+1}=u_j^n-\left[\sigma+q\left(\frac{\psi_{j+\frac{1}{2}}}{\eta_{j+\frac{1}{2}}}-\psi_{j-\frac{1}{2}}\right)\right]\Delta u_{j-1}^n$ 均具有二阶精度。若同时满足 $\frac{\sigma}{2+\sigma}\leqslant\eta$，则满足 TVD 的要求。

现设 $\psi(\eta)=1-\delta(\eta)+\eta\delta(\eta)$，$0\leqslant\delta(\eta)\leqslant1$，当 $\delta=0,\psi=1$；$\delta=1,\psi=\eta$，只要再满足 $\frac{\sigma}{2+\sigma}\leqslant\eta\leqslant1$，则可满足 TVD。下面各式就是二阶精度的 TVD 格式。

三、各种 TVD 格式

1. Van Leer 二阶 TVD 格式

$$\delta=\frac{1}{1+|\eta|},\psi=1-\frac{1}{1+|\eta|}+\frac{\eta}{1+|\eta|}=1+\frac{\eta-1}{1+|\eta|}$$

2. Superbee 二阶 TVD 格式

$$\psi_{Roe}=\max\{0,\min(2\eta,1),\min(\eta,2)\}$$

这是诸多限制器中压缩最严重的，故称 Superbee。

3. $\psi_{yee}(\eta)=\min\bmod(\eta,1)$

还有其他多种当 ψ 与 η 的关系式，满足 TVD 的要求，如 $\psi_{co}(\eta)=\max\{0,\min(\eta,\phi)\}$，其中 $1\leqslant\phi\leqslant2$。

四、注意事项

(1) 前面的 TVD 格式是由 $u_t+au_x=0$，即线性方程导出的。由于 $\psi(\eta)$ 与 $\Delta u_{j\pm1}$ 相联系，TVD 差分格式是非线性的。

(2) 在守恒情况下，初值问题还可能出现二阶精度 TVD 的非弱解（稀疏激波解），TVD 无法解决此问题，q,ψ 必须按熵条件作修正处理。

五、结果举例

Van Leer TVD 与 Lax-Wendroff 格式比较

$$\psi=1+\frac{\eta-1}{1+|\eta|}=1-\frac{1-\eta}{1+|\eta|}$$

$$\begin{aligned}
u_j^{n+1}&=u_j^n-\sigma\Delta u_{j-1}^n-q\left(\frac{\psi_{j+\frac{1}{2}}}{q_{j+\frac{1}{2}}}-\psi_{j-\frac{1}{2}}\right)\Delta u_{j-1}^n\\
&=u_j^n-\sigma\Delta u_{j-1}^n-q(\psi_{j+\frac{1}{2}}\Delta u_j^n-\psi_{j-\frac{1}{2}}\Delta u_{j-1}^n)\\
&=u_j^n-\sigma\Delta u_{j-1}^n-q\left[\left(1-\frac{1-\eta_{j+\frac{1}{2}}}{1+|\eta_{j+\frac{1}{2}}|}\right)\Delta u_j-\left(1-\frac{1-\eta_{j-\frac{1}{2}}}{1+|\eta_{j-\frac{1}{2}}|}\right)\Delta u_{j-1}\right]\\
&=u_j^n-\sigma(u_j^n-u_{j-1}^n)-\frac{\sigma}{2}(1-\sigma)(u_{j+1}^n-2u_j^n+u_{j-1}^n)\\
&\quad+\frac{\sigma}{2}(1-\sigma)\cdot\left[\beta_{j+\frac{1}{2}}(u_{j+1}^n-u_j^n)-\beta_{j-\frac{1}{2}}(u_j^n-u_{j-1}^n)\right]
\end{aligned}\tag{4.14}$$

式中 $\beta_{j+\frac{1}{2}}=\dfrac{1-\eta_{j+\frac{1}{2}}}{1+|\eta_{j+\frac{1}{2}}|}, \beta_{j-\frac{1}{2}}=\dfrac{1-\eta_{j-\frac{1}{2}}}{1+|\eta_{j-\frac{1}{2}}|}$。

可见 Van Leer TVD 格式在 Lax-Wendroff 格式基础上又加上正耗散，即(4.14)等号右侧第 4 项，表明 Lax-Wendroff 在一阶迎风基础上多减了粘性，即负耗散加得太多了，要增加些正耗散既保证二阶精度又保证 TVD。

4.1.2 由二阶中心差分格式构成 TVD 格式

用上述原理构成的二阶精度 TVD 格式具有迎风性质，捕捉激波能力强。下面通过二阶中心差分格式构造 TVD 格式。

仍分析微分方程(1.58)式，将其写成 Lax-Wendroff 格式为：

$$u_j^{n+1}=u_j^n-\frac{\sigma}{2}(u_{j+1}^n-u_{j-1}^n)+\frac{\sigma^2}{2}(u_{j+1}^n-2u_j^n+u_{j-1}^n) \tag{2.29}$$

此时(2.29)式可理解为中心差分格式(无耗散)由于计算不稳定加上了正耗散而得，此格式二阶精度不变且稳定，但不是 TVD，原因与中心差分无耗散相比，其正耗散加得还不够，现写成用一般 $Q_{j\pm\frac{1}{2}}$ 来调节的格式：

$$\begin{aligned}u_j^{n+1}=&u_j^n+\frac{\sigma}{2}(1+\sigma)\Delta u_{j-1}^n-\frac{\sigma}{2}(1-\sigma)\Delta u_j^n-\frac{|\sigma|}{2}(1-|\sigma|)(1-Q_{j-\frac{1}{2}})\Delta u_{j-1}^n\\&+\frac{|\sigma|}{2}(1-|\sigma|)(1-Q_{j+\frac{1}{2}})\Delta u_j^n\end{aligned} \tag{4.15}$$

当 $Q_{j\pm\frac{1}{2}}=0$ 时，(4.15)变成

$$u_j^{n+1}=u_j^n-\sigma(u_j^n-u_{j-1}^n)$$

这是一阶迎风格式，正耗散最大。

当 $Q_{j\pm\frac{1}{2}}=1$ 时，(4.15)变成：

$$u_j^{n+1}=u_j^n-\sigma(u_j^n-u_{j-1}^n)-\frac{\sigma}{2}(1-\sigma)(u_{j+1}^n-2u_j^n+u_{j-1}^n)$$

这是 Lax-Wendroff 格式，正耗散最小。调节 $0<Q_{j\pm\frac{1}{2}}<1$ 可得 TVD 格式。常采用的 Q 的表达式为：

(1) $Q=\min\bmod(1,\eta^-)+\min\bmod(1,\eta^+)-1$

$$\eta_{j+1/2}^-=\frac{\Delta u_{j-1}}{\Delta u_j}, \eta_{j+1/2}^+=\frac{\Delta u_{j+1}}{\Delta u_j}, \text{保证 } 0<Q<1 \tag{4.16}$$

(2) $Q=\min\bmod(1,r^-,r^+)$

$$r_{j-1/2}^+=\frac{\Delta u_{j-2}}{\Delta u_{j-1}}, r_{j-1/2}^-=\frac{\Delta u_j}{\Delta u_{j-1}} \tag{4.17}$$

(3)
$$Q=\frac{r^-+|r^-|}{1+r^-}+\frac{r^++|r^+|}{1+r^+}-1 \tag{4.18}$$

因此，Lax-Wendroff 格式为中心差分格式，为了保持二阶精度且具有 TVD 性质，可在原 Lax-Wendroff 格式的基础上，加上可调节的正的人工粘性，得下式右侧的第二、三

项，形成中心型二阶 TVD 差分格式：

$$u_j^{n+1} = (2.29)\text{式右侧} + \Delta u_j\left(1 - Q_{j+\frac{1}{2}} + \frac{\sigma^2}{2}\right) - \Delta u_{j-1}\left(1 - Q_{j-\frac{1}{2}} + \frac{\sigma^2}{2}\right) \quad (4.19)$$

4.1.3 一阶偏微分方程组的 TVD 格式

对于方程组：

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = \frac{\partial U}{\partial t} + A\frac{\partial U}{\partial x} = 0 \quad (2.34)$$

式中 $U=(u_1,u_2,\cdots,u_m)^T$，$F=(F_1,F_2,\cdots,F_m)^T$，$A=\frac{\partial F}{\partial u}$，是 Jacobian 系数矩阵，通常是 $m\times m$ 阶。

(2.34)是双曲型方程组，A 矩阵有 m 个线性独立的实特征值。根据第二章所述，(2.34)可写成：

$$\frac{\partial \omega}{\partial t} + \Lambda\frac{\partial \omega}{\partial x} = 0 \quad (4.20)$$

其中 $\omega=P^{-1}U$，$\Lambda=P^{-1}AP=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \ddots & \vdots\\ \vdots & 0 & \ddots & 0\\ 0 & \cdots & 0 & \lambda_m\end{bmatrix}$。$P$ 是由右特征向量组成的矩阵，P^{-1}为其逆矩阵。于是(4.20)可写成：

$$\begin{cases}\dfrac{\partial \omega_1}{\partial t} + \lambda_1\dfrac{\partial \omega_1}{\partial x} = 0\\ \dfrac{\partial \omega_2}{\partial t} + \lambda_2\dfrac{\partial \omega_2}{\partial x} = 0\\ \quad\vdots\\ \dfrac{\partial \omega_m}{\partial t} + \lambda_m\dfrac{\partial \omega_m}{\partial x} = 0\end{cases} \quad (4.21)$$

于是可将前面得到的由(1.58)所导出的 TVD 格式分别用于(4.21)的每个方程上，只是每一个方程的 a 分别由 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 代替，最后的解为 $U=P\omega$。

事实上，Harten，Van Leer 和 Osher 等人均通过

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \frac{\partial u}{\partial t} + a(u)\frac{\partial u}{\partial t} = 0 \quad (4.22)$$

这种非线性的守恒型标量方程来构造 TVD 格式，其原理与前面的分析类似，这里不再详述。正如前面所讲的，守恒型的方程在求解有间断的可压流场时有其优越性，因此构成的 TVD 格式具有广泛的应用。

4.1.4 基本无振荡的高阶格式(ENO)

前面讲述的差分格式，包括 TVD 格式，至多只能达到二阶精度，在进行工程上复杂外形流场计算时，在计算机目前能够承受的条件下，往往在某些方向(如 x、y)，网格数满足不了把物理粘性项的贡献正确反映出来的要求，而是被格式的数值截断误差所掩盖。若

采用高阶精度的格式，则对各向网格数的要求可放松。

另外，目前在应用计算空气动力学中，一定网格下，物面压力分布可以用二阶精度格式算得很好，符合工程需要，但对粘性摩擦阻力系数与热流率等要求物面物理量计算准确后才能确定的量，常常误差较大，无法满足工程需要；同时，对一些微小扰动引起的影响，如边界层转捩分析中，要求对小扰动源及噪声等都应分辨清楚。以上这些都希望有高阶精度的差分格式的出现。

1987 年，Harten 与 Osher 等人提出了 3 阶或更高精度的基本无振荡格式(Essentially Non-Oscillatory ，简称 ENO)，其特点是守恒，基本上具有 TVD 性质，具有一致的高阶精度的差分格式。ENO 与 TVD 相比，除了精度提高，总变差由递减变为有界外，主要是在极值点处保持空间的高精度，而二阶 TVD 格式在极值点处退化为一阶精度。

下面对 ENO 的构成作一简单介绍。

$$\text{微分方程}:\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}=0 \tag{4.22}$$

$$\text{初值问题}:u(0,x)=u_0(x)$$

一、在网格单元建立平均解

$$\bar{u}(t,x)=\frac{1}{h}\int_{-\frac{h}{2}}^{\frac{h}{2}}u(t,x+s)\mathrm{d}s \tag{4.23}$$

$u(t,x)$是(4.22)的解。后面构造的格式均是指由 $s=-\frac{h}{2}$ 到 $s=\frac{h}{2}$ 的单元平均值 $\bar{u}$.

二、函数重构

在 $t=t_n,x=x_j$ 的网格点处变量 $\bar{u}$ 为 $\bar{u}_j^n$。由 $\bar{u}_j^n$ 重构一个函数 $R(x,\bar{u}^n)$，并由初值求解守恒方程：

$$\omega_t+f(\omega)_x=0,t\in[t^n,t^{n+1}] \tag{4.24}$$

$$\omega(t^n,x)=R(x,\bar{u}^n),x\in(-\infty,\infty)$$

三、由(4.24)所得的解 $\omega(t^n+\tau,x)$再进行网格平均

$$\bar{u}_j^{n+1}=\frac{1}{h}\int_{-\frac{h}{2}}^{\frac{h}{2}}\omega(t^n+\tau,x_j+s)\mathrm{d}s,j=-\infty,\cdots\cdots,+\infty$$

这样构成的解尚需满足一系列条件，详见有关资料。我们用五阶精度的加权的 ENO 即 WENO，计算了 $M_\infty=2.0$ 时绕平板层流边界层与激波干扰引起的板面粘性摩擦应力，并与实验结果和二阶 NND 格式的计算结果进行了比较，WENO 计算所取的网格数和其他条件与 NND 格式一样。图 4.1 是无分离时的结果，图 4.2 是层流分离时的结果。由图可见，五阶 WENO 格式计算出的粘性应力系数与实验值十分接近，效果较好。ENO 格式及其应用目前正在发展。

4.2 加快收敛速率的一些措施

在定常或弱非定常流场的数值计算中，为了提高计算效率，近年来进行了 LU-SGS

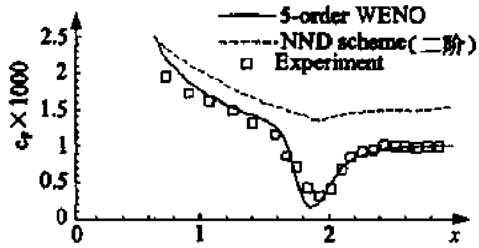


图 4.1　无分离时 WENO 和 NND 格式所得摩擦应力分布(网格 81×81)

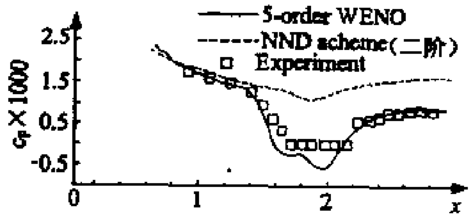


图 4.2　有分离时 WENO 和 NND 格式所得摩擦应力分布(网格 81×81)

与 LU-SSOR 的方法及预处理方法的研究和使用。这对于计算流体力学的工程应用有很大的推动作用，现择要介绍如下。

4.2.1　定常流场或弱非定常流的隐式 LU 法

由于定常流场计算的时间相关法中，物理量随时间的变化无明显的物理含义，时间步长 $k(\Delta t)$ 只起迭代参数的作用。以(3.15)式 $\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0$ 为例，最终收敛后方程转化为：

$$\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0$$

此时 $U^{n+1}\Rightarrow U^n$，$\delta U^n\Rightarrow 0$，所以在隐式计算，如(3.31a)式 $(I+\frac{k}{2}\frac{\partial A}{\partial x})^n\overline{U^{n+1}}$ 的计算中，可以大大简化，即使时间精度不高，但不影响最终的 U^{n+1} 的精度，而空间精度主要由显式部分来确定的，与 Δt 无关。

一、LU-SSOR 法(Lower-Upper Symmetric Successive Over Relaxation)

1. Euler 方程的 LU-SSOR 方法

微分方程(Euler 方程)：

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \tag{3.15}$$

由(1.1)知：

$$U=(\rho,\rho u,\rho v,\rho E)^T$$
$$F=[\rho u,p+\rho u^2,\rho uv,u(\rho E+p)]^T$$
$$G=[\rho v,\rho uv,p+\rho v^2,v(\rho E+p)]^T$$
$$p=(\gamma-1)\rho(E-\frac{u^2+v^2}{2})$$

Beam-Warming 法：

$$U^{n+1}=U^n-\beta k\left[\frac{\partial F^n}{\partial x}+\frac{\partial F^{n+1}}{\partial x}+\frac{\partial G^n}{\partial y}+\frac{\partial G^{n+1}}{\partial y}\right]$$
$$=U^n-\beta k[D_xF^n+D_xF^{n+1}+D_yG^n+D_yG^{n+1}]$$

当 $\beta=\frac{1}{2}$ 时即(3.27)式。如果不进行 $F^n=A^nU^n$，$F^{n+1}=A^nU^{n+1}$ 则上式具有二阶时间精度。

但在数值求解时如：

$$(I+\frac{k}{2}D_yB)U^{n+1}=(I+\frac{k}{2}\frac{\partial B}{\partial y})U^{n+1}=\overline{U^{n+1}} \tag{3.32b}$$

当对$\frac{\partial}{\partial y}$进行中心差分时，要解$U^{n+1}_{j,m}$，$U^{n+1}_{j,m+1}$，$U^{n+1}_{j,m-1}$三对角块矩阵求逆，工作量非常大，现在进行一些数值简化工作：

令 $$\delta U=U^{n+1}-U^n$$

则 $$F^{n+1}=F(U^{n+1})=F^n+A^n\delta U \tag{4.25a}$$

$$G^{n+1}=G(U^{n+1})=G^n+B^n\delta U \tag{4.25b}$$

代入前面的 Beam-Warming 式中得：

$$[I+\beta k(D_xA+D_yB)]\delta U+\beta kR=0,\quad R=D_xF^n+D_yG^n \tag{4.26}$$

若取 $\beta=1$ 即为全隐式；若令 $k\to\infty$，则有：

$$(D_xA+D_yB)\delta U=-R \tag{4.27}$$

若采用迎风形式，则更有利于捕捉激波。对(4.27)式进行矢通量分裂：

$$(D_x^-A^-+D_x^-A^++D_y^+B^-+D_y^-B^+)\delta U=-R \tag{4.28}$$

一阶迎风格式空间精度不高，但不影响最终结果，因为$\delta U_{t\to\infty}=0$，于是(4.28)可改写为：

$$\begin{aligned}&A^+_{j,m}\delta U_{j,m}-A^+_{j-1,m}\delta U_{j-1,m}+A^-_{j+1,m}\delta U_{j+1,m}-A^-_{j,m}\delta U_{j,m}\\&+B^+_{j,m}\delta U_{j,m}-B^+_{j,m-1}\delta U_{j,m-1}+B^-_{j,m+1}\delta U_{j,m+1}-B^-_{j,m}\delta U_{j,m}=-R_{j,m}\end{aligned} \tag{4.28a}$$

对上式进行两步超松弛迭代求解。

第一步：使 j,m 由小到大，$j-1,j,m-1,m$ 等处的物理量已知，求 $j+1,m+1$ 处物理量的值：

$$(A^+_{j,m}-A^-_{j,m})\delta U^*_{j,m}+A^-_{j+1,m}\delta U^*_{j+1,m}+(B^+_{j,m}+B^-_{j,m})\delta U^*_{j,m}+B^-_{j,m+1}\delta U^*_{j,m+1}=-R'_{j,m} \tag{4.29}$$

x 方向步进时只有 $\delta U^*_{j+1,m}$未知，y 方向步进时只有 $\delta U^*_{j,m+1}$未知，故可直接求解，不必三对角求逆。

第二步：

$$\begin{aligned}&(A^+_{j,m}-A^-_{j,m})\delta U_{j,m}-A^+_{j-1,m}\delta U_{j-1,m}+(B^+_{j,m}-B^-_{j,m})\delta U_{j,m}\\&-B^+_{j,m-1}\delta U_{j,m-1}+A^-_{j+1,m}\delta U^*_{j+1,m}+B^-_{j,m+1}\delta U^*_{j,m+1}=-R_{j,m}\end{aligned} \tag{4.30}$$

返回扫描，此时 $j+1,m,j,m+1$ 均为已知值，只有 $\delta U_{j-1,m}$与 $\delta U_{j,m-1}$未知，因此也不必三对角求逆，可以大大节省计算时间。

进一步简化运算：

$$\begin{aligned}A^\pm&=\frac{1}{2}(A\pm r_AI)\\B^\pm&=\frac{1}{2}(B\pm r_BI)\\r_A&\geqslant\rho_A=\max(|\lambda_A|)\\r_B&\geqslant\rho_B=\max(|\lambda_B|)\end{aligned} \tag{4.31}$$

由(4.30)减(4.29)得：

$$(A_{j,m}^{+}-A_{j,m}^{-})\delta U_{j,m}+(B_{j,m}^{+}-B_{j,m}^{-})\delta U_{j,m}-A_{j-1,m}^{+}\delta U_{j-1,m}-B_{j,m-1}^{+}\delta U_{j,m-1}$$
$$=(A_{j,m}^{+}-A_{j,m}^{-})\delta U_{j,m}^{*}+(B_{j,m}^{+}-B_{j,m}^{-})\delta U_{j,m}^{*}$$

上式也可写为：

$$(D_x^- A^+ + D_y^- B^+ - A^- - B^-)\delta U = (A^+ + B^+ - A^- - B^-)\delta U^* \tag{4.32}$$

$D_x^{\pm}$ 等为一阶迎风差分算子。由(4.31)得：

$$A^+ - A^- = r_A I,\quad B^+ - B^- = r_B I \tag{4.33}$$

由(4.29)得：

$$\delta U^* = (D_x^+ A^- + D_y^+ B^- + A^+ + B^+)^{-1}(-R),$$

代入(4.32)得：

$$\begin{aligned}&(D_x^- A^+ + D_y^- B^+ - A^- - B^-)(D_x^+ A^- + D_y^+ B^- + A^+ + B^+)\delta U\\&= -(D_x F^n + D_y G^n)(r_A + r_B)I\end{aligned} \tag{4.34}$$

(4.34)的 ADI：

$$(D_x^- A^+ + D_y^- B^+ - A^- - B^-)\delta \underline{U} = -(D_x F^n + D_y G^n)(r_A + r_B)I \tag{4.35a}$$

$$(D_x^+ A^- + D_y^+ B^- + A^+ + B^+)\delta U = \delta \underline{U} \tag{4.35b}$$

方程左侧的 D_x, D_y 均为一阶迎风差分，格式的空间精度由 $R=(D_xF^n+D_yG^n)$ 来决定，当 $\delta U\to 0, n\to\infty$ 时，若用前面的二步方法求解，未知量的块矩阵求解是比较简便的直接求解，不必求逆。因为是定常流场，时间精度不影响结果。

2. N. S. 方程的 LU-SSOR 方法

对 N. S. 方程前面在讲多维流稳定性准则时也讲过，方程(4.36)是其模型方程：

$$\frac{\partial U}{\partial t} + A\frac{\partial U}{\partial x} = \upsilon\frac{\partial^2 U}{\partial x^2} \tag{4.36}$$

由 3.2 节知其稳定性判断准则为：

$$(\Delta t)^{-1} = (\Delta t_x)^{-1} + (\Delta t_y)^{-1} \geqslant \frac{2|\rho_A|}{\Delta x} + \frac{4\upsilon}{\Delta x^2}$$

$$\Delta t \leqslant \frac{\Delta x^2}{4\upsilon + 2|\rho_A|\Delta x} \tag{3.14}$$

对空气的流动，$\bar{\gamma}=\upsilon=\frac{\mu}{\rho}$ 的值为：

$$\upsilon = 1.57\times 10^{-5}\mathrm{m^2/s}(300\mathrm{K\ air}),$$
$$\upsilon = 11.2\times 10^{-5}\mathrm{m^2/s}(1000\mathrm{K\ air})$$
$$\rho_A \sim (u+a), M=0.5, \rho_A = u+a \Rightarrow 1.5a$$
$$\Delta t \leqslant \frac{\Delta x^2}{4\times 1.6\times 10^{-5} + 2\times 1.5\times 340\cdot\Delta x}$$

相比之下，分母第二项，即无粘对流项起主要作用。

隐式的目的是为提高 Δt，不受 CFL 稳定性之限制，而粘性的存在对 Δt 几乎无影响，主要是对流项起关键作用，于是粘性部分只取显式，隐式部分采用与上述 Euler 方程相同的(4.34)式。

$$(D_x^- A^+ + D_y^- B^+ - A^- - B^-)(D_x^+ A^- + D_y^+ B^- + B^+ + A^+)\delta U$$

$$
\begin{aligned}
&= -(r_A + r_B)I\{[D_x(F_I^n - F_V^n) + D_y(G_I^n - G_V^n)] \\
&\quad - q_{x2}D_x^2U + q_{x4}D_x^4U - q_{y2}D_y^2U + q_{y4}D_y^4U\}
\end{aligned} \tag{4.37}
$$

(4.37)式中等号右侧 F_I,G_I 为无粘通量,F_V,G_V 为粘性通量。因右侧的 D_x,D_y 用二阶中心差分,所以后面加上了人工粘性项;若等号右侧 D_x,D_y 用 TVD、NND 或二阶迎风离散,则不必这样加入人工粘性。

二、LU-SGS 法(Lower-Upper Symmetric Gauss-Seidel Relaxation)

与 LU-SSOR 之区别在于未取 $\Delta t \to \infty$,而用牛顿迭代:

$$
\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0 \tag{1.1}
$$

由(4.26)得:

$$
[I/\Delta t + \beta(D_xA + D_yB)]\delta U + \beta R = 0 \tag{4.38}
$$

$\beta=1$ 时为全隐式,则有:

$$
(I/\Delta t + D_xA + D_yB)\delta U + R = 0 \tag{4.39}
$$

矢通量分裂后一阶迎风为:

$$
\left(\frac{I}{\Delta t} + D_x^- A^+ + D_y^- B^+ + D_x^+ A^- + D_y^+ B^-\right)\delta U = -R
$$

由(4.28a)及(4.31)得:

$$
\left[\frac{I}{\Delta t} + (r_A + r_B)I + (A_{j+1,m}^- + B_{j,m+1}^-) - (A_{j-1,m}^+ + B_{j,m-1}^+)\right]\delta U = -R \tag{4.40}
$$

注意:(4.40)中 δU 的下标随 A,B 的下标而变动,对无 A,B 的项则为 $\delta U_{j,m}$。令 $(\frac{1}{\Delta t}+r_A+r_B)I=d$,则(4.40)为:

$$
[d + (A^- + B^-) - (A^+ + B^+)]\delta U = -R
$$

令:$d-(A^+ + B^+)=L \quad d+B^- + A^- = u$,上式变为:

$$
[Ld^{-1}u]\delta U = -R \tag{4.41}
$$

具体解(4.41)的过程为:

$$
L\delta U^* = -R
$$

$$
(d^{-1}u)\delta U = \delta U^*
$$

解出 δU 后,$U^{n+1}=U^n+\delta U$。

上述两种做法只适用于定常流与弱非定常流,原因是在时间上与原方程不相容,时间精度低。

4.2.2 标量化对角矩阵隐式求解

对非定常流要求与原方程相容且时间精度较高,则不能用上述方法。现由 Beam-Warming 隐格式出发进行说明。(3.33)可写成:

$$
\left(I + \frac{k}{2}\frac{\partial A\cdot}{\partial x}\right)\left(I + \frac{k}{2}\frac{\partial B\cdot}{\partial y}\right)\delta U = -RHS \tag{4.42}
$$

其中 $RHS=k\left(\frac{\partial A^n\cdot}{\partial x}+\frac{\partial B^n\cdot}{\partial y}\right)U^n$。

现用 $A=P_A\Lambda_AP_A^{-1}$,$B=P_B\Lambda_BP_B^{-1}$,P_A,P_B 是分别由 A,B 矩阵的右特征向量组成的矩

阵代入(4.42)得：

$$P_A(I+\frac{k}{2}\frac{\partial\Lambda_A}{\partial x})P_A^{-1}P_B(I+\frac{k}{2}\frac{\partial\Lambda_B}{\partial y})P_B\delta U=-RHS$$

令 $P_A^{-1}P_B=N$，上式变为：

$$P_A(I+\frac{k}{2}\frac{\partial\Lambda_A}{\partial x})N(I+\frac{k}{2}\frac{\partial\Lambda_B}{\partial y})P_B^{-1}\delta U=-RHS$$

用 ADI 方法求解时，可表述为：

$$\begin{aligned}&P_A(I+\frac{k}{2}\frac{\partial\Lambda_A}{\partial x})\delta\bar{U}=-RHS\\&(I+\frac{k}{2}\frac{\partial\Lambda_A}{\partial x})\delta\bar{U}=-R_1RHS=-P_A^{-1}RHS\end{aligned}\tag{4.42a}$$

$$\begin{aligned}&N(I+\frac{k}{2}\frac{\partial\Lambda_B}{\partial y})P_B^{-1}\delta U=\delta\bar{U}\\&(I+\frac{k}{2}\frac{\partial\Lambda_B}{\partial y})\delta U=\delta\bar{U}^*=N^{-1}\delta\bar{U}P_B\end{aligned}\tag{4.42b}$$

P_A,P_B,N 均用已知的第 n 层值，$\frac{\partial\Lambda_A}{\partial x},\frac{\partial\Lambda_B}{\partial y}$用中心差分，即使是三对角块矩阵求解，由于只是三对角求逆，计算量大大减少，但二阶与四阶人工粘性不能忘记，因为在离散 RHS，$\frac{\partial\Lambda_A}{\partial x},\frac{\partial\Lambda_B}{\partial y}$时用了中心差分。这种做法在时间精度上没有损失。

4.2.3 预处理法

在计算流体力学中，时间相关法(Time-Dependent Method)被广泛用于可压流的计算，但在小马赫数及不可压流中，此方法由于方程中对流项的刚性问题而收敛很慢，表现较差。我们试图用解可压流的时间相关法来高效地求解小马赫数流场和不可压流场，这就是这里将要介绍的预处理法(Preconditioning)。预处理方法对于求解既含有低速流，又含有高速流的大迎角流场非常有效，且可用于求解含有大的加热量的低速流场。这也是利用可压缩流数值方法解低速不可压流的一个途径。当然，它也只能用于定常流场和弱非定常流的求解。

一、不可压流场的人工压缩法

定常不可压流场的微分方程组为：

$$\left.\begin{aligned}&\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0\\&u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial x}=\frac{1}{\mathrm{Re}}(\frac{\partial^2u}{\partial x^2}+\frac{\partial^2u}{\partial y^2})\\&u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial y}=\frac{1}{\mathrm{Re}}(\frac{\partial^2v}{\partial x^2}+\frac{\partial^2v}{\partial y^2})\end{aligned}\right\}\tag{4.43}$$

方程具有椭圆型方程的性质，但又有不确定性，不能利用初值问题的技术进行研究，若用 Time-Dependent 法，则方程改写为：

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0,\frac{\partial\rho}{\partial t}\equiv0$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial x}=\frac{1}{\mathrm{Re}}(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}) \tag{4.43}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial y}=\frac{1}{\mathrm{Re}}(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2})$$

方程类型仍具有不确定性，在不可压情况下具有双曲一椭圆的性质，在 Time-Dependent 法的前提下，即允许使用与时间相关方程不相容而与定常方程相容的预处理法，即所谓人工压缩法。

此时方程改为：

$$\frac{\partial p}{\partial t}+a^2(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y})=0$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial x}=\frac{1}{\mathrm{Re}}(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}) \tag{4.44}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{1}{\rho}\frac{\partial p}{\partial y}=\frac{1}{\mathrm{Re}}(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2})$$

其中 $\rho=\text{constant}$。

在非定常情况下解无物理意义，因为与原方程不相容，尤其是质量方程，但在定常状态下与原方程相容，把质量方程视为“可压缩连续方程”，因此得人工可压缩性的名称。

a 可称人工声速，$p=a^2\rho$，事实上计算中 ρ 并不显示，于是 a 与 Δt 可视为迭代参数，Δt 是由显式计算的稳定性限制所定，事实上 a 不能随便给出，其限制后面再讲。

可用显式 Dufort-Frankel 及其他显式方法包括 MacCormack 法，当然也可用隐式 Beam-Warming 近似因子法求解，现介绍之。将(4.44)写成列矢量形式：

$$\frac{\partial Q}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}-\frac{1}{\mathrm{Re}}D\nabla^2 Q=0 \tag{4.45}$$

其中

$$Q=(p,u,v)^T,F=(a^2u,p+u^2,uv)^T$$

$$G=(a^2v,uv,p+v^2)^T,D=\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

令 $A=\frac{\partial F}{\partial Q},B=\frac{\partial G}{\partial Q}$，则：

$$A=\begin{bmatrix}0&a^2&0\\1&2u&0\\0&v&u\end{bmatrix},B=\begin{bmatrix}0&0&a^2\\0&v&u\\1&0&2v\end{bmatrix}$$

$$F=AQ-uDQ=\begin{bmatrix}0&a^2&0\\1&2u&0\\0&v&u\end{bmatrix}\begin{bmatrix}p\\u\\v\end{bmatrix}-u\begin{bmatrix}0&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}p\\u\\v\end{bmatrix}$$

$$=\begin{bmatrix}a^2\\p+2u^2\\2uv\end{bmatrix}-u\begin{bmatrix}0\\u\\v\end{bmatrix}=\begin{bmatrix}a^2u\\p+u^2\\uv\end{bmatrix} \tag{4.46a}$$

$$G=BQ-vDQ \tag{4.46b}$$

用 Crank-Nicolson 格式离散(4.45),

$$\frac{\partial Q^n}{\Delta t}+\frac{L_x}{2}(F^n+F^{n+1})+\frac{L_y}{2}(G^n+G^{n+1})=\frac{D}{2\mathrm{Re}}(L_{xx}+L_{yy})(Q^n+Q^{n+1}) \quad (4.47)$$

其中 $L_x=\frac{\bar{\delta}_x}{2\Delta x}$和 $L_{xx}=\frac{\delta_x^2}{\Delta x^2}$均为中心差分算子,$\delta Q^n=Q^{n+1}-Q^n$。

由 $F^{n+1}\approx F^n+A^n\delta Q^n=F^n+\frac{\partial F}{\partial Q}\Delta Q$, $G^{n+1}=G^n+B\delta Q^n$ 得:

$$\begin{cases}\left\{I+\frac{\Delta t}{2}\left[L_xA^n+L_yB^n-\frac{D}{\mathrm{Re}}(L_{xx}+L_{yy})\right]\right\}\delta Q^n=RHS \\ RHS=\Delta t\left[(D/\mathrm{Re})(L_{xx}+L_{yy})Q^n-L_xF^n-L_yG^n\right]\end{cases} \quad (4.48)$$

用 ADI 方法可表述为:

$$\left[I+\frac{k}{2}\left(L_xA^n-\frac{D}{\mathrm{Re}}L_{xx}\right)-\varepsilon_2\Delta x^2L_{xx}\right]\delta Q^*=RHS-\varepsilon_4\left[(\nabla_x\Delta_x)^2+(\nabla_y\Delta_y)^2\right]Q^n \quad (4.49a)$$

$$\left[I+\frac{k}{2}\left(L_yB^n-\frac{D}{\mathrm{Re}}L_{yy}\right)-\varepsilon_2\Delta y^2L_{xx}\right]\delta Q^n=\delta Q^* \quad (4.49b)$$

其中$(\nabla_x\Delta_x)^2Q^n=Q^n_{j-2,k}-4Q^n_{j-1,k}+6Q^n_{j,k}-4Q^n_{j+1,k}+Q^n_{j+2,k}$,$\varepsilon_2$,$\varepsilon_4$ 要求小于$\frac{1}{\mathrm{Re}}$以免数值耗散掩盖物理耗散,本式的四阶导数用以克服非线性的不稳定性,隐式二阶光滑用以平衡显式二阶光滑项,以避免在非定常的计算中在接近定常解时收敛太慢。

二、预处理法(Preconditioning)

1. 概述

预处理法最初由上述人工压缩法发展而来,但效果更好,应用范围也更广。预处理法的实质是在时间相关法的时间导数项前乘一个预处理矩阵,因而改变了控制方程的特征值,使可压流控制方程在低马赫数流场中,特征值的刚性问题得以解决,这个方法只能用于定常流场的求解,因为加上预处理矩阵后,现方程与非定常流控制方程不相容,但相容于定常流的控制方程,无论是 Euler 方程和 N.S. 方程均如此。这个方法收敛快,精度不降低,可用到全马赫数流场的求解,把此思想延伸,也可用于解决化学非平衡流等求解中的刚性问题。

二维非定常可压流 N.S. 方程是(1.1)式

$$\frac{\partial U}{\partial t}+\frac{\partial(F_I-F_V)}{\partial x}+\frac{\partial(G_I-G_V)}{\partial y}=0$$

前已说明,当用显式差分格式求解时,稳定性限制主要由方程的无粘部分系数矩阵的最大特征值的绝对值所确定,粘性部分影响不大。如以一维非定常无粘流控制方程(Euler 方程)为例,其系数矩阵的特征值由(1.50)式得

$$\lambda_{1,2,3}=[u+a,u,u-a]^T$$

当来流马赫数很小时,三个特征值差别很大,如 $M=\frac{u}{a}=0.01$,则特征值间相差 100 倍,引入条件数 K

$$K=\frac{|\lambda|\max}{|\lambda|\min}=1+\frac{1}{M} \quad (4.50)$$

在显格式计算中 K 的增大，反映流场中信息传播的差异，即在 Δt 时间间隔中，最大与最小波速传播的距离相差 K 倍，使计算中收敛很慢。采用预处理方法，就是使 K 减小，在同一 Δt 下，几种波的传播距离类同，从而加快了收敛。而控制方程是时间相关法所对应的可压流方程，因此可以用时间相关法解可压流问题的全部技巧来解决不可压流场问题。

2. N. S. 方程的预处理法

在(1.1)式的时间导数项前乘一个预处理矩阵 Γ

$$\Gamma \frac{\partial Q}{\partial t} + \frac{\partial (F_I - F_V)}{\partial x} + \frac{\partial (G_I - G_V)}{\partial y} = 0 \tag{4.51}$$

(4.51)与(1.1)相比，除了乘上矩阵 Γ 外，时间导数项的变量 $U=(\rho,\rho u,\rho v,E_1)^T$ 也改成了原始变量 $Q=(\rho,u,v,T)^T$，式中

$$\Gamma = \begin{pmatrix} \frac{1}{T} & 0 & 0 & \frac{-\rho}{T} \\ \frac{u}{T} & \rho & 0 & \frac{-\rho u}{T} \\ \frac{v}{T} & 0 & \rho & \frac{-\rho v}{T} \\ C_v + \frac{q}{T} & \rho u & \rho v & \frac{-\rho q}{T} \end{pmatrix} \tag{4.52}$$

其中 $q=\frac{1}{2}(u^2+v^2)$。

现分析(4.51)在一维非定常情况下无粘部分系数矩阵的特征值，令 $A=\frac{\partial F_I}{\partial Q}$，然后求 $\Gamma^{-1}A$ 的特征值得：

$$A = \begin{pmatrix} \frac{u}{RT} & \rho & \frac{-\rho u}{T} \\ \frac{u^2}{RT}+1 & 2\rho u & \frac{-\rho u^2}{T} \\ \frac{uH}{RT} & \rho(H+u^2) & \frac{-\rho q u}{T} \end{pmatrix}, \Gamma^{-1} = \begin{pmatrix} \frac{q}{C_v} & \frac{-u}{C_v} & \frac{1}{C_v} \\ \frac{-u}{\rho} & \frac{1}{\rho} & 0 \\ \frac{q-C_vT}{\rho C_v} & \frac{-u}{\rho C_v} & \frac{1}{\rho C_v} \end{pmatrix} \tag{4.53}$$

式中 C_v 是等容比热，R 是气体常数，由此求得 $\Gamma^{-1}A$ 的特征值：

$$\lambda_1 = u, \quad \lambda_{2,3} = \frac{1}{2}\left[(1+\frac{1}{R})u \pm \sqrt{(1-\frac{1}{R})^2u^2 + 4\gamma T}\right] \tag{4.54}$$

可以证明，当来流马赫数 $M=\frac{u}{a}$ 很小，$\frac{1}{R}$ 是 $O(M^2)$ 量级，从而保证条件数 K 为 $O(1)$ 量级，因此加快了收敛。

预处理矩阵 Γ 的构造是本法成败的关键，对一个方程，可构成多种预处理矩阵，因此确定预处理矩阵的设计准则十分重要。首先，Γ 应能消除刚性并具有正定性和对称性，这是求得适定和稳定的解所必须的。其次特征值的优化是重要的准则，另外在小马赫数时，对流动角不敏感和没有平行的特征矢量等准则也要注意。

利用可压流方程的预处理方法求解不可压流问题，效果不错。计算时可采用显式多步 Runge-Kutta 法或隐式 ADI 法及 LU-SGS 法(见 4.2.1 节)，在解定常粘性湍流 N. S. 方

程时，常用 LU-SGS 法。

4.3 结构网格与非结构网格

对应用于工程上的计算流体力学来说，由于飞机、导弹、船舶、车辆等物体外形相当复杂，正确高效的生成网格是计算成败的关键问题之一。目前生成的网格主要有两类，即结构网格与非结构网格。结构网格的节点以阵列形式排列，与计算机语言自然匹配，便于矩阵演算与操作，本书 3.6 节所述的网格设计都属于结构网格的生成方法。应当指出，除了 3.6 节所介绍的椭圆型偏微分方程网格生成外，现常用的还有双曲型方程网格生成法。

非结构网格对网格的节点没有结构性的限制，其节点与单元的分布具有任意性，因此具有优越的几何灵活性，生成网格的质量相对较高，近年来受到国内外计算流体力学工作者的高度重视。非结构网格的缺点是相对于结构网格需要更多的计算机内存及 CPU 时间；在计算流体力学中较成熟的加快收敛的方法如多重网格法等不能在非结构网格中应用，在粘性复杂流场计算中非结构网格也存在着一些尚需解决的问题。

本节主要介绍适用于复杂外形的结构网格中的分区网格生成法和非结构网格生成法。

4.3.1 分区网格生成

分区网格生成的基本思想是把复杂外形构成的流场划分为若干个子区域，在每个子区域中独立生成适合于其外形的网格，然后把各子区域网格连接起来，使复杂外形的网格问题得以解决，这是结构网格用于复杂外形的有效方法。采用分区网格生成法有如下优点：

(1) 采用分区网格生成能较好地解决用单区网格生成无法解决或效果不佳的复杂外形流场的计算问题。

(2) 分区网格生成后，不同区域允许采用不同的网格系统或用不同的差分格式求解。

(3) 在分区网格生成后的求解过程中，任何时刻，只有一个子区的数据存放于计算机缓冲区内，从而放宽了对计算机存贮器的要求。

(4) 采用分区网格计算，对于在多处理机计算机上实现多机并行计算提高并行计算的效率是有效的。

分区网格系统中各相邻子区有重叠部分的称为覆盖或重叠网格(Overlapping grids)，和各相邻子区无重叠的称拼接网格或对接网格(Patched grids)，我们将分别讨论。

一、分区拼接网格

分区拼接网格的生成大致归纳为下述四个步骤：

(1) 根据几何外形与流动特点进行分区，并对每区中网格拓扑进行分析。如图 4.3 所示，沿气流方向廓线在 *EFGH* 处间断，如用代数网格生成方法，按整个流场生成整体网格，以 *AEFGHB* 为一个边界，以 *AD*、*DC*、*CB* 为另三条边界，四边值问题生成网格，有如下的缺点：网格与物面不正交；网格扭曲、尖点多；网格走向与流动方向不一致，这样生成的网格将严重影响计算的精度及效果，如按分区拼接的思想，将计算域分成三个独立的四

边形(见图 4.3),上述缺点就迎刃而解了。

(2) 生成表面网格。这是对于工程中复杂外形必须采用方法。

(3) 交界面网格的生成。对于一个三维流场,相邻区域间的公交面,通常是一个空间曲面,可使用各种方法在空间曲面上生成网格。如方程网格生成法或代数网格生成法。

(4) 当表面与交界面上的网格生成后,最终可用方程网格生成法或多边值问题代数网格生成法生成空间网格。

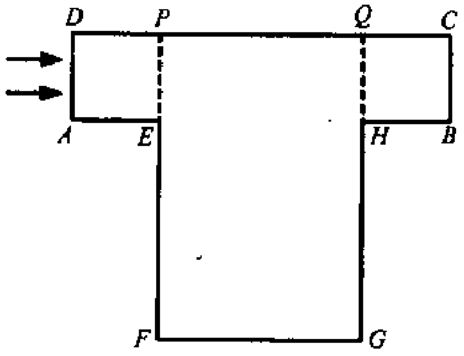


图 4.3 分区图形

在分区网格生成后,进行数值计算时可用如下步骤进行:

① 每一个子区内解控制方程组求得内点上守恒变量值。

② 应用区域边界格式,在区域边界上进行守恒性处理,以保证区域边界上通量守恒。

③ 通过插值求得相邻子区域边界上的守恒边界值。

④ 在相邻子区域内,求解控制方程组,得到其内点上的变量值。以此类推,可求得全流场的变量值。

上述拼接网格法的区域边界信息量交换很少,处理比较简单,在可压流跨区域激波捕捉过程中,如保证区域边界变量传递守恒,则激波在跨区域处是连续的。

二、重叠网格生成

上述拼接网格生成法有一个基本的要求,即各子区域间要有一个共同的边界,这使其使用范围受到一定的限制。而重叠网格中,各子区域间不需有共同边界,只需包含有主网格区内的一个重叠区域进行不同网格域间的信息交换。所以重叠网格法又称嵌入式分区网格(Chimera grid)。

在划分网格时,为了使分割后各子区外形简单,往往将有重叠网格部分的子区从主域上去掉,于是重叠网格就变成在包含它的主网格上的一个孔(Hole),孔内的网格单元不参加主域的流场计算,但它是各区域变量传递的工具。下面就重叠网格的生成中网格等级、孔的形成及网格间的通讯分别作一说明。

1. *网格等级*

重叠网格法生成网格时,每个子区是单独生成的,因此,首先要确定不同网格的等级关系,即分清各子区网格是并列的、包含的、还是交叉的。在图 4.4 中示意了等级关系,图中的某一网格 $G_{l,j}$,l 代表其级别,$j=1,2,\cdots$代表 l 级的若干个网格,我们规定 $l+1$ 级网格完全包含在 l 级网格内。

2. *孔的形成*

重叠网格区域中的网格关系很复杂,例如,主网格 $G_{l,j}$的节点可能包含在 $G_{l+1,j}$的固壁内部,但这些点不属于计算域,应除去;在信息交换时不必要把 $G_{l,j}$与 $G_{l+1,j}$所有共同点的值均更新,这样做工作量太大而无必要。为了提高插值效率,我们将 $G_{l,j}$包含在 $G_{l+1,j}$内的某些点从主计算域中除去,就会在 $G_{l+1,j}$内引入了一个孔(Hole),见图 4.5,孔的大小可

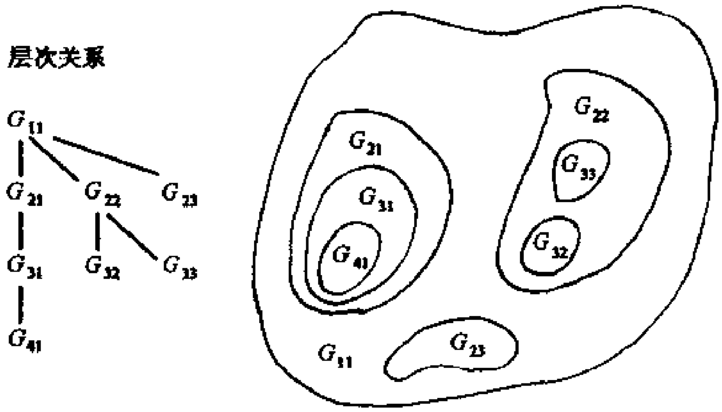


图 4.4　嵌入式网格等级结构

按计算情况调节，但重叠区域也要有足够的共同节点以传递信息。

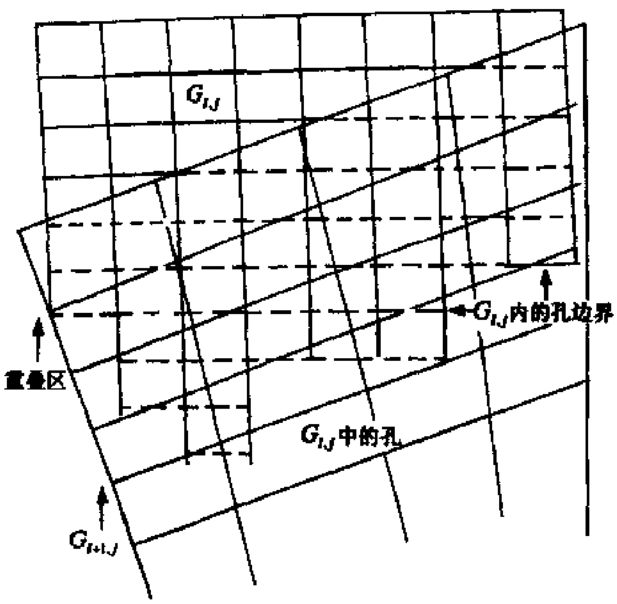


图 4.5　重叠区的网格

具体计算中，孔的生成可按下列方法进行：

在 $G_{i+1,j}$中选定曲线 C 作为孔边界，对 $G_{i,j}$中每一网格点，设状态参数 IB 用以标识孔的信息：$\begin{cases} IB=1 & \text{该点不在孔内} \\ IB=0 & \text{该点在孔内} \end{cases}$，图 4.6 对 $G_{i,j}$网点状态的确定作了说明。

下面是确定 $G_{i,j}$网格点状态的步骤(见图 4.6)：

(1) 求出 C 曲线上每个节点上外法线矢量 $\boldsymbol{N}$，并在曲线内设一个临时原点 P_0；

(2) 求曲线上各点到原点的最大距离 R_{max}；

(3) 对 $G_{i,j}$中每一个网格点 P 求出线段$\overline{PP_0}$；

(4) 若$\overline{PP_0}>R_{max}$，则 P 点在曲线 C 之外，若$\overline{PP_0}<R_{max}$，则 P 有可能在 C 内。确定曲线 C 上离 P 最近之点 P_c，及 $\boldsymbol{R}_p$ 是 P_c 到 P 的矢量，算出 $\boldsymbol{R}_p \cdot \boldsymbol{N}$，$\boldsymbol{N}$ 是 P 点的外法线矢量。

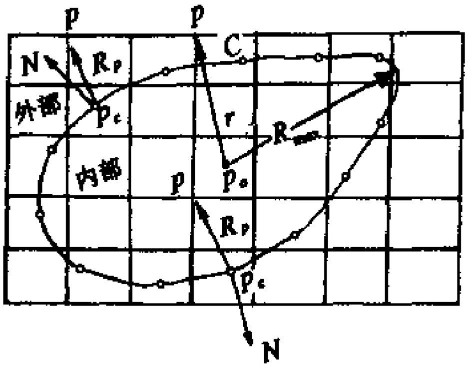


图 4.6　孔的生成

(5) 当 $\boldsymbol{R}_p \cdot \boldsymbol{N}<0$，则 P 点在曲线 C 内，此时设定 $IB=0$，孔的边界可取与 $IB=0$ 相邻的 $IB=1$ 的点，选定了孔的边界后，可由此边界点从 $G_{l+1,j}$中通过插值获取信息。

3. 网格间的信息传递

网格间的信息传递是通过插值来实现的。$G_{l+1,j}$的外边界上的变量值由 $G_{l,j}$中插值而得，而 $G_{l,j}$孔边界点上的变量由包含它的 $G_{l+1,j}$中获取，插值对重叠网格至关重要，应尽量减少插值误差。通常可采用双线性插值法，其效率较高，操作简单，具体作法可参阅例如 Mastin C. W. 等人的论文。

4.3.2　非结构网格

由于结构网格在拓扑结构上相当于矩形域内的均匀网格，其节点定义在每一层的网格线上，且每一层上节点数都是相等的，这样使复杂外形的贴体网格生成比较困难。非结构网格没有规则的拓扑结构，也没有"层"这一概念，其网格节点分布是随意的，因此具有灵活性。20 世纪 80 年代末 90 年代初以来，非结构网格获得迅速发展。

非结构网格的特点有下列四个：

① 适于处理复杂外形的边界。非结构网格大都采用最简的几何外形，三角形（二维）与四面体（三维）单元，易于剖分不规则计算域，由此种单元建立起来的离散方程相对简单。

② 根据流场性质自由合理安排网格节点，如流场中物理量变化剧烈处，网格安排就密，物理量变化平缓的地方网格点可较稀。

③ 可作自适应处理。

④ 网格生成省时。

目前，非结构网格生成方法主要是两大类，即 Delaunay 的三角化法和阵面推进法。Delaunay 三角化方法可能构成区域外的或和边界相交的单元，使流场的边界完整性遭到破坏，但它生成三角单元质量高，网格生成效率相对较高。阵面推进法是网格与节点同时生成的非结构网格生成法，其三角元是从边界开始形成，然后向内（外）推进，内部网格节点在阵面推进中自动引进，最终使三角形网格覆盖整个计算域。本书以二维（三角元）阵面推进法为例简要介绍网格生成的过程。其主要步骤是：生成背景网格；对内外边界进行剖分形成初始阵面；引入新节点推进阵面，生成网格；对生成的网格进行光滑与优化。

(1) 背景网格。指的是一组覆盖整个计算域的疏网格，其节点处存贮着网格生成所需的控制参数，主要是单元的基本长度 l、网格的伸展比 δ 和伸展方向 ω，为简单起见，我们这里只用其基本长度 l。背景网格不是最终网格的组成部分，不需与边界保持一致，但此

时计算域内任一点的控制参数都可用其所在的背景网格三个顶点的参数线性插值得到，保证了网格单元分布的连续性。最基本的背景网格是人工输入数据形成，现在新的改进的措施是用 Delaunay 法自动生成。

(2) 初始阵面的形成。对二维网格来说，阵面是由边(矢量)组成，初始阵面可用背景网格提供的控制参数，对物面边界和远场边界分别进行剖分得到。控制体外边界的边(矢量)按逆时针方向组织；内边界也是一个个有向闭环，但每一个内边界的边按逆时针方向排列，这些边最终在三角形化结构中成为三角形的边。三角化的最初阵面一般来说是所有内边界的总和。

(3) 阵面推进，生成网格。初始阵面形成后，可以由背景网格提供的控制参数引入内部结点，生成网格。这个过程包括连接新节点或已存在的阵面节点形成三角单元并对阵面进行重新定义，其具体步骤可参考 Parikh 等有关资料。

(4) 对网格进行优化。为进一步提高网格的质量，还需对上面生成的网格进行优化，其最常用的方法是用 Laplacian 迭代法：

$$x_j^{n+1} = x_j^n + \frac{\omega}{m}\sum_{k=1}^{m}(x_k^n - x_j^n)$$

$$y_j^{n+1} = y_j^n + \frac{\omega}{m}\sum_{k=1}^{m}(y_k^n - y_j^n)$$

其中 ω 是松弛因子，求和符号取遍以节点 j 为端点的三角形边。这种方法将内部节点移到该节点周围三角元所组成的多边形形心处，可对网格起光顺作用。

4.3.3 结构/非结构混合网格

如前所述，非结构网格相比于结构网格存在着内存要求大、CPU 时间长、不能使用结构网格中有效地加快收敛的措施和粘流计算中非结构网格的生成尚需进一步研究等问题，于是提出了结构网格与非结构网格混合网格的问题。

矩形网格是结构网格中最简单的一种，不必进行 Jacobian 矩阵计算，具有比贴体网格计算更简便快捷的优点，但不易处理复杂边界。因此，最简单的混合网格是由近物面的非结构网格及远物面的矩形网格的混合，这类混合网格在无粘流计算时是很有效的。

也有用分区重叠结构网格，而在其交界处的孔，由大量非结构网格来填充，从而实现相邻两网格间的通量守恒，于是形成了所谓拉链网格(Zipper grid)和龙形网格(Dragon grid)。

第五章 流体力学中的有限元素法

在流体力学，尤其在高速流体力学或气体动力学，有限差分方法占着统治地位，但是它对不规则边界条件的处理任意性小，处理复杂，缺乏严格的论证，常常需要数值试验来确定方法的优劣，即它是后验的。目前在数值解抛物型与椭圆型方程中，有限元法渐渐显示出其优越性，且渐渐向求解双曲型方程渗透。

因为有限差分法发展早，运用广，早有成套的软件，要舍弃这些软件另搞新的有限元法，成本上与效率上都不合算。这可能也是有限元法在流体力学上还未占优势的原因，但有人预料有限元法将来有可能在纯数值计算上与有限差分方法平分秋色，例如边界上用边界有限元，内点用有限差分处理，效果就比较好。近来，由于非结构网格的发展，使有限元法在计算流体力学中的应用越来越广泛。

本章主要就流体力学有限元法求解中应用较广的迦辽金(Galerkin)加权余量法作了介绍。在分析了其机理后，叙述了一维、二维有限元法的求解要素。二维问题主要介绍椭圆型方程的求解，其中也简略分析了变分法的基本原理及其与迦疗金法之关联。本章末就精度问题、边界条件处理问题及流体力学控制方程求解问题作了简单介绍。本章只是流体力学中有限元法的基础。

5.1 加权余量法(WRM)

加权余量法(Weighted Residual Method)主要用于将偏微分方程积分进行数值求解。

一、基本思路

设微分方程为：

$$\bar{L}(u) = p \tag{5.1}$$

式中 $u = u(x)$ 是方程的解，$x \in V$ 是解域 V 上的空间坐标(即 x,y,z)，$\bar{L}$ 是微分算子。例如：$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} - v\frac{\partial^2 u}{\partial x^2} = 0 \Rightarrow \bar{L} = \frac{\partial}{\partial t} + a\frac{\partial}{\partial x} - v\frac{\partial^2}{\partial x^2}$，$p = 0$。$p$ 是空间矢量的已知函数。

边界条件为：

$$L(u) = g(x), x \in S \tag{5.2}$$

其中 g 是已知函数，S 是解域 V 的边界面，L 是已知算子。例如，当 $\frac{\partial u}{\partial n} = g$ 时 $L = \frac{\partial}{\partial n}$。

二、求解方法

1. *试验函数(或近似函数)*

$$\bar{u} = \sum_{k=1}^{N} \alpha_k \phi_k \tag{5.3}$$

其中 $\bar{u}$ 是近似解，α_k 是待定的未知参数，ϕ_k 则是完备序列的线性无关函数，称基函数(通常是空间坐标的显函数)，如 $u = f(x) \Rightarrow \bar{u} = A + Bx + Cx^2 + Dx^3$。

2. 余量

因为 $\bar{u} \neq u$，故有

$$\varepsilon = \bar{L}(\bar{u}) - p \neq 0 \tag{5.4}$$

称为余量，在整个求解域 V 内 ε 分布有任意性。

α_k 的选取对余量的影响也有任意性，我们现在要想办法使在平均意义下强迫 $\varepsilon \to 0$，由此可确定各个 α_k，则问题即可解（ϕ_k 是按问题的性质与精度的要求事先选定的）。

三、数学表达

$$\langle \varepsilon, w_i \rangle = \int_V \varepsilon(x) w_i(x) \mathrm{d}V = 0 \tag{5.5}$$

式中 w_i 称为加权函数。为了确定 α_k 使 ε 最小，我们选取 N 个权函数 $w_i (i = 1, 2, \cdots, N)$，并令 ε 与 w_i 正交，即其内积为 0，这样就可列出 N 个方程，解出 $\alpha_k (k = 1, 2, \cdots, N)$，其物理意义即是在 N 个位置上以某种形式强迫 ε 与 w_i 正交，于是问题可解。

小结：微分方程 $\bar{L}(u) = p \to \bar{u} = \sum_{k=1}^{N} \alpha_k \phi_k \to \phi_k$ 选定（如何选定以后讲）$\to \varepsilon = \bar{L}(\bar{u}) - p \to \langle \varepsilon, w_i \rangle = 0 \to \alpha_k \to \bar{u}$ 求得近似解。当然要求满足定解条件；这是与有限差分法不同的求解思路。

本章主要解椭圆型方程，且主要讲基本概念与线性方程的求解，即最基本的基础知识。

5.1.1 基本概念

一、内积

两个函数 $\phi_1(x), \phi_2(x)$ 的内积可表示为：

$$\langle \phi_1, \phi_2 \rangle = \int_{x_1}^{x_2} \phi_1(x) \phi_2(x) \mathrm{d}x \tag{5.6}$$

就实函数而言，内积的性质如下：

$$\langle \phi_1, \phi_2 \rangle = \langle \phi_2, \phi_1 \rangle$$
$$a\langle \phi_1, \phi_2 \rangle = \langle a\phi_2, \phi_1 \rangle$$
$$\langle \phi_1, \phi_2 + \phi_3 \rangle = \langle \phi_1, \phi_2 \rangle + \langle \phi_1, \phi_3 \rangle$$
$$\langle \phi_1, \phi_1 \rangle > 0, \text{当 } \phi_1 \neq 0$$
$$\langle \phi_1, \phi_1 \rangle = 0, \text{当 } \phi_1 = 0$$

其尺度标志是本身内积的方根，即 $\| \phi \| = \sqrt{\langle \phi, \phi \rangle}$，称为范数(Norm)。内积的这种性质使我们想起了这与两个 N 维矢量的标量积(点积)相类似，其范数犹如矢量的长度。

二、完备序列与基函数

定义　有一 N 个函数系列 $\phi_1(x), \phi_2(x), \cdots, \phi_N(x)$，如果 $\sum_1^N \alpha_i \phi_i = 0$ 只有当所有的 $\alpha_i = 0$ 时才成立，则上述函数系列是线性无关的。

如果给定一个任意函数 u，总可找到数 N 及一组常系数 α_i 使 $\| u - \sum_1^N \alpha_i \phi_i \| < \delta$，$\delta$ 是

任意小量，则称上述线性无关序列 $\alpha_i\phi_i$ 是完备的，函数系列 ϕ_i 称为基函数，α_i 称为付立叶系数。

规范化的基函数如互相正交，则

$$\langle\phi_i,\phi_j\rangle=\begin{cases}0,i\neq j\\1,i=j\end{cases}$$

证明 将 ϕ_i,ϕ_j 视为多维向量。

$\langle\phi_i,\phi_j\rangle=\|\phi_i\|\ \|\phi_j\|\cos(\phi_i,\phi_j)$。规范化的基函数 $\|\phi_i\|=1,\|\phi_j\|=1$。

ϕ_i,ϕ_j 正交 $\Rightarrow$ 夹角 $(\phi_i,\phi_j)=\dfrac{\pi}{2}$，从而 $\langle\phi_i,\phi_j\rangle=0$，且 $\langle\phi_i,\phi_i\rangle=1$。证毕。

5.1.2 一维有限元 WRM 法中权函数 w_i 的选择

常用的是迦辽金法，但也介绍最小二乘法与配置法。

一、配置法(Collocation)

用 Dirac 函数(又称脉冲函数)作为权函数。

定义 $\langle\varepsilon,\Delta(x-x_i)\rangle=\varepsilon|_{x=x_i}=0,\ i=1,2,\cdots,N$，其含义是在固定的配置点 x_i 上将余量 ε 取为 0，其它非固定点上 ε 未加限制，以此来保证整个积分区 ε 趋于零。

例如：

$$\frac{d^2u}{dx^2}+u+x=0=\bar{L}(u)-p \tag{5.7}$$

边界条件：$\begin{cases}x=0, & u=0\\x=1, & u=0\end{cases}$，解域 $0\leqslant x\leqslant 1$。

设

$$\bar{u}=\sum_{k=1}^{2}\alpha_k\phi_k=\alpha_1[x(1-x)]+\alpha_2[x^2(1-x)] \tag{5.8}$$

用前而介绍的方法求 α_1,α_2。

(1) 满足边界条件：检查 $x=0,\bar{u}=0$；$x=1,\bar{u}=0$，合理。

(2) $$\varepsilon=\bar{L}(\bar{u})-p=\frac{d^2\bar{u}}{dx^2}+\bar{u}+x=x+(x-2-x^2)\alpha_1+(2-6x+x^2-x^3)\alpha_2 \tag{5.9}$$

(3) 选取 $x_{i1}=0.25,x_{i2}=0.5$ 作为配置点。

$$\langle\varepsilon,\Delta(x-x_i)\rangle=\varepsilon|_{x=x_i}=0$$

$$\begin{cases}\varepsilon(x=0.25)=0\\\varepsilon(x=0.5)=0\end{cases}\Rightarrow\begin{cases}\dfrac{29}{16}\alpha_1-\dfrac{35}{64}\alpha_2=\dfrac{1}{4}\\[2ex]\dfrac{7}{4}\alpha_1-\dfrac{7}{8}\alpha_2=\dfrac{1}{2}\end{cases}$$

解出：$\alpha_1=\dfrac{6}{31}$，$\alpha_2=\dfrac{40}{217}$，从而 $\bar{u}_C=\dfrac{x(1-x)}{217}(42+40x)$，它与精确解 $u_*=\dfrac{\sin x}{\sin 1}-x$ 比较，最大相对误差为 3.5%。

二、最小二乘法(使余量本身的内积最小，方程同上)

$$\varepsilon=\bar{L}(\bar{u})-p,\ F=\langle\varepsilon,\varepsilon\rangle$$

$$\frac{\partial F}{\partial \alpha_i}=0,\ i=1,2,\cdots,N$$

展开得：

$$\frac{\partial}{\partial \alpha_i}[\langle[\bar{L}(\bar{u})-p],[\bar{L}(\bar{u})-p]\rangle]=\frac{\partial}{\partial \alpha_i}[\langle\bar{L}(\bar{u}),\bar{L}(\bar{u})\rangle-2\langle\bar{L}(\bar{u}),p\rangle+\langle p,p\rangle]=0$$

列出 N 个方程可解出 α_i 值。

当 $\bar{L}$ 为线性算子时，则可有下述结果：

$$\frac{\partial}{\partial \alpha_i}\langle\bar{L}(\sum\alpha_k\phi_k),\bar{L}(\sum\alpha_k\phi_k)\rangle-2\frac{\partial}{\partial \alpha_i}\langle\bar{L}(\sum\alpha_k\phi_k),p\rangle=0$$

$$2\langle\bar{L}(\phi_i),\bar{L}(\bar{u})\rangle-2\langle\bar{L}(\phi_i),p\rangle=0$$

即：

$$\langle\bar{L}(\phi_i),(\bar{L}(\bar{u})-p)\rangle=0 \quad 或 \quad \langle\varepsilon,\bar{L}(\phi_i)\rangle=0$$

可见最小二乘法的权函数 $w_i=\bar{L}(\phi_i)$，即将基函数代入微分方程算子中去，解上述(5.7)式得

$$\bar{L}(\phi_1)=\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}+1\right)[x(1-x)]=x-2-x^2,$$

$$\bar{L}(\phi_2)=\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}+1\right)[x^2(1-x)]=2-6x+x^2-x^3$$

其中 ϕ_1 和 ϕ_1 可由(5.8)求得：$\phi_1=x(1-x),\phi_2=x^2(1-x)$。于是

$$\int_0^1\varepsilon(x-x^2-2)\mathrm{d}x=0$$

$$\int_0^1\varepsilon(2-6x+x^2-x^3)=0$$

其中的 ε 见(5.9)式。

求出 $\bar{u}_L=x(1-x)(0.1875+0.1695x)$，，最大相对误差为 2.4%。

三、迦辽金法

将基函数 ϕ_i 作为权函数：

$$\langle\varepsilon,\phi_i\rangle=0$$

$$\int_0^1\left\{\bar{L}\left(\sum_{k=1}^{2}\alpha_k\phi_k\right)-p\right\}\phi_i\mathrm{d}x=0$$

$$\int_0^1\varepsilon x(1-x)\mathrm{d}x=0$$

$$\int_0^1\varepsilon x^2(1-x)\mathrm{d}x=0$$

解得 $\alpha_1=\dfrac{71}{369}$，$\alpha_2=\dfrac{7}{41}$，$\bar{u}_G=\dfrac{x(1-x)}{369}(71+63x)$，最大相对误差为 0.4%。

x	u_*	$\bar{u}_G$	$\bar{u}_L$	$\bar{u}_C$
0.25	0.04401	0.04408	0.04310	0.04493
0.5	0.06974	0.06944	0.06806	0.07143
0.75	0.06005	0.06008	0.05899	0.06221

可以看出迦辽金法精度最高。并非所有问题的求解结果都如此，但迦辽金法确实便于流体

力学运用且可适用于非线性方程，精度也比较高，因此应用最广。

常用的迦辽金法的等价公式为：

$$\delta\bar{u} = \sum_{i=1}^{N}\delta\alpha_i\phi_i = (\delta\alpha_i)\phi_i,\ i = 1,2,\cdots,N$$

这里相同指标 i 应求和，$\delta\alpha_i$ 是 α_i 的任意增量，$\delta\bar{u}$ 是 $\bar{u}$ 的任意增量。

一维：$\langle\epsilon,\phi_i\rangle = \int\epsilon\phi_i\mathrm{d}x = \int\epsilon\delta\alpha_i\phi_i\mathrm{d}x = \int\epsilon\delta\bar{u}\mathrm{d}x = \int[L(\bar{u} - p]\delta\bar{u}\mathrm{d}x = 0$ (5.10)

三维：$\int_V\epsilon\phi_i\mathrm{d}V = 0 \to \int_V\epsilon\delta\alpha_i\phi_i\mathrm{d}V = 0 \to \int_V\epsilon\delta\bar{u}\mathrm{d}V = 0 = \int_V[L(\bar{u}) - p]\delta\bar{u}\mathrm{d}V$

$i = 1,2,\cdots,N$ (5.11)

二维公式也可类似导出。

四、应用举例

考虑 Poisson 方程：$\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} - c = 0$

边界条件：$\begin{cases} x = 0,a \\ y = 0,b \end{cases} \Rightarrow u = 0$

最简单的近似形式 $\bar{u} = \alpha_1\phi_1 = \alpha_1(x-a)xy(y-b)$ 显然满足边界条件。

由(5.11)式得：$\int_0^a\int_0^b\left[\left(\dfrac{\partial^2\bar{u}}{\partial x^2} + \dfrac{\partial^2\bar{u}}{\partial y^2} - c\right)\right]\delta\bar{u}\mathrm{d}x\mathrm{d}y = 0$。其中 $\delta\bar{u} = \delta\alpha_1(x-a)(y-b)xy$ 是 α 任意变化时 $\bar{u}$ 的增量。

$$\frac{\partial^2\bar{u}}{\partial x^2} = 2\alpha_1 y(y-b),\ \frac{\partial^2\bar{u}}{\partial y^2} = 2\alpha_1 x(x-a)$$

代入上式，积分得：$\alpha_1 = \dfrac{5}{2}c\dfrac{1}{a^2+b^2}$

从而 $\bar{u}_c\left(x = \dfrac{a}{2},\ y = \dfrac{b}{2}\right) = -\dfrac{5}{32}c\dfrac{a^2b^2}{a^2+b^2}$

若 $a = b$，则 $\bar{u}_c = -0.078125a^2c$，而在 $x = \dfrac{a}{2}$，$y = \dfrac{b}{2}$ 时的精确解为 $-0.07357a^2c$，这里的误差是由一阶近似引起的，而且在其他位置误差更大。

5.1.3 基函数及其分析

一、基函数 ϕ 的选取

现在讨论一般的一维有限元法。把解区划分为若干块(i 块)。

设 $\bar{u} = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots = f(x) = y$ 为多项级数，在各分界处有 $y_0, y_1, y_2, \cdots, y_5$，则应当有：

$$\left.\begin{aligned} y_0 &= \sum_{i=1}^{5}\alpha_i x_0^i \\ y_1 &= \sum_{i=1}^{5}\alpha_i x_1^i \\ &\vdots \\ y_5 &= \sum_{i=1}^{5}\alpha_i x_5^i \end{aligned}\right\}$$

6 个方程，当 $y_i(i = 0,1,2,\cdots,5)$ 已知时，解 6 个未知数 $\alpha_i(i = 0,1,2,$

$\cdots,5)$，问题可解，$\bar{u}$ 可求。

总之，若用 N 次多项式时有 $N+1$ 个节点，解 $N+1$ 个试验函数值。首先，$\bar{u}$ 并不一定是幂级数，但总能写成在节点上的某个试验函数，且解在此点上的未知值 α_i 而构成试验函数表达式。我们想到了 Lagrange 插值函数：

$$l_i(x_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{5.12}$$

$$\bar{u} = \sum_{i=0}^{N} \bar{u}_i l_i(x_j) = \sum_{k=0}^{N} \alpha_k \phi_k \tag{5.13}$$

于是待求值 α_k，即 $\bar{u}$ 就是各节点上的应有值。基函数 $\phi_k = l_i(x_j)$ 保证在节点上均为 $\bar{u}$ 的当地应有值，其它地方 $l_i(x_j) = 0$。

$$l_i(x) \triangleq \frac{(x-x_0)(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_N)}{(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_N)} \tag{5.14}$$

当 $x = x_i$ 时，$l_i(x) = 1$；当 $x \neq x_i$ 时，$l_i(x) = 0$。

可见选用 Lagrange 插值多项式为基函数可保证在计算时各节点适定。待求函数在节点上为其应有值，这正是我们所希望的，应保证待求函数在各节点处存在。

下面举例进行说明。

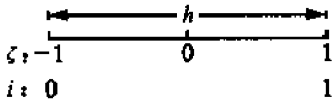


图 5.1

1. 一维线性分布元素

一个元素两个节点：

$$\bar{u}|_{x=0} = \bar{u}_0, \bar{u}|_{x=1} = \bar{u}_1$$

规范化，用 h 除 x，并将原点置于元素的中间，$\bar{x} = \dfrac{x}{h}$，$\zeta = 2\bar{x} - 1$，$\begin{cases} \bar{x} = 0, \zeta = -1 \\ \bar{x} = 0.5, \zeta = 0 \\ \bar{x} = 1, \zeta = 1 \end{cases}$。以后均用无因次量，用 ζ 是为了规范化，并与后面的二维元素相衔接。

若 $\bar{u}$ 是线性分布，$\bar{u} = \alpha_0 + \alpha_1\zeta$，代入得：

$$\left.\begin{array}{l} \bar{u}_0|_{\zeta=-1} = \alpha_0 - \alpha_1 \\ \bar{u}_1|_{\zeta=1} = \alpha_0 + \alpha_1 \end{array}\right\} \Rightarrow \alpha_0 = \frac{\bar{u}_0 + \bar{u}_1}{2}, \ \alpha_1 = \frac{\bar{u}_1 - \bar{u}_0}{2}$$

所以

$$\bar{u} = \frac{\bar{u}_0 + \bar{u}_1}{2} + \frac{\bar{u}_1 - \bar{u}_0}{2}\zeta = \sum_{i=0}^{1} \bar{u}_i \phi_i$$

$$\phi_0 = \frac{1-\zeta}{2}, \ \phi_1 = \frac{1+\zeta}{2} \tag{5.15}$$

于是基函数求出，并且显然满足边界条件。

事实上可用 Lagrange 多项式求出。由前面的(5.13)和(5.14)式得：

$$\bar{u} = \bar{u}_0 l_0 + \bar{u}_1 l_1 = \sum_{i=0}^{1} \bar{u}_i l_i$$

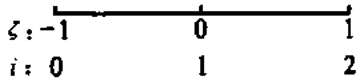


图 5.2

$$l_0(\zeta)=\frac{(\zeta-\zeta_1)}{(\zeta_0-\zeta_1)}=\frac{1-\zeta}{2}$$

$$l_1(\zeta)=\frac{(\zeta-\zeta_0)}{(\zeta_1-\zeta_0)}=\frac{1+\zeta}{2}$$

与上面的结果完全一样。

2. *变量二次变化的基函数*

一个元素上变量二次变化，见图 5.2。

节点应选 3 个：0,1,2。对应的 ζ=−1,0,1。代入(5.14)得：

$$\begin{aligned} l_0 &= \frac{\zeta(1-\zeta)}{2} \\ l_1 &= 1-\zeta^2 \\ l_2 &= \frac{\zeta(1-\zeta)}{2} \\ \bar{u} &= \sum \bar{u}_k l_k = \sum \bar{u}_k \phi_k \end{aligned} \tag{5.16}$$

并保证满足边界条件。

如果取近似解是变量三次函数，则取四个点，例如 $\zeta=-1,-\frac{1}{3},\frac{1}{3},1$。照样可列出，依次类推。总之用 Lagrange 插值多项式作基函数，保证节点处物理量 $\bar{u}$ 连续(即为当地值)，元素中变量的变化可按线性、二次函数甚至更高阶函数分布来描述。阶次高，精度就越高，但计算量更大。

二、Hermite 插值函数作为基函数

令 n 是 WRM 积分式中泛函中的最高阶，那么基函数的导数须在元素间或元素内连续到 $n-1$ 阶(以后要证明)。通常为了使高阶微分方程求解时基函数不太复杂，可采用降阶的办法。

Laplace 方程在写成 WRM 积分形式后，可由分部积分的办法降阶，即：

$$\iint_S\left(\frac{\partial^2\bar{u}}{\partial x^2}+\frac{\partial^2\bar{u}}{\partial y^2}\right)\delta\bar{u}\,\mathrm{d}x\mathrm{d}y=0\Rightarrow\iint_S\varepsilon\phi_i\,\mathrm{d}x\mathrm{d}y=0$$

$$\Rightarrow\frac{\delta}{2}\iint\left[\left(\frac{\partial\bar{u}}{\partial x}\right)^2+\left(\frac{\partial\bar{u}}{\partial y}\right)^2\right]\mathrm{d}x\mathrm{d}y+\left(\frac{\partial\bar{u}}{\partial x}\delta\bar{u}+\frac{\partial\bar{u}}{\partial y}\delta\bar{u}\right)_{BV}=0$$

(其中下标 BV 表示边界值)

由于 $\frac{\partial^2\bar{u}}{\partial x^2}\delta\bar{u}=\frac{\partial}{\partial x}\left(\frac{\partial\bar{u}}{\partial x}\partial\bar{u}\right)-\frac{\partial\bar{u}}{\partial x}\delta\frac{\partial\bar{u}}{\partial x}=\frac{\partial}{\partial x}\left(\frac{\partial\bar{u}}{\partial x}\delta\bar{u}\right)-\delta\frac{1}{2}\left(\frac{\partial\bar{u}}{\partial x}\right)^2$，$\frac{\partial^2\bar{u}}{\partial y^2}\delta\bar{u}$ 也可类似导出，可见积分式内最高阶导数由 $n=2$ 降到 $n=1$。此时，只要求元素边界上有 $n-1=0$，即零阶连续，即只要求函数值连续就行。Lagrange 插值函数满足了此要求，而在某元素上取自变量 ζ 的二次三次变化只是在元素内的分布状态，不影响元素边界上函数的光滑程度。但对于双调和方程 $\nabla^4 u=0$，分部积分只能降阶到“2”阶，$n=2$，$n-1=1$，要求元素间一阶

导数连续，此时用 Lagrange 插值函数无法满足光滑性的要求，用 Hermite 插值函数可达到此要求：

$$\bar{u} = y(x)$$
$$H^0(x_i) = y(x_i) = \bar{u}(x_i) \tag{5.17a}$$

$$\left(\frac{\mathrm{d}H}{\mathrm{d}x_j}\right)_i = H'(x_i) = y'(x_i) = \bar{u}'(x_i) \tag{5.17b}$$

$$\bar{u}(x) = \sum_{i=1}^{N} y(x_i)H_i^0(x) + \sum_{i=1}^{N} y'(x_i)H'_i(x) \tag{5.18}$$

它与 Lagrange 插值多项式的试验函数有两个差异：一是每一节点处有两个未知量，$y'(x_i)$ 与 $y(x_i)$，而 L 式只有一个量；二是基函数本身的阶次，H 式要求为 2^m+1 次（$m=n-1$ 为连续导数的阶次），而 L 式要求为 m 次；$m=1$ 的 H 式为三阶多项式，而 $m=0$ 的 L 式为二阶多项式。

例如：一维元素两端为节点，变量的一阶导数连续，先进行无因次化，原点不移动，由(5.18)得：

$$\bar{u}(\zeta) = \sum_{1}^{2} y(\zeta_i)H_i^0(\zeta) + \sum_{1}^{2} y'(\zeta_i)H_i^1(\zeta)$$

要求

$$H_i^0(\zeta_j) = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}H_i^1(\zeta_j) = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$$

一般形式：

$$H(\zeta) = \sum_{i=1}^{N} a_i[1+2(\zeta_i-\zeta)b_i]y(\zeta_i) + \sum_{i=1}^{N} a_i(\zeta-\zeta_i)y'(\zeta_i)$$
$$= \sum_{i=1}^{N} y(\zeta_i)H_i^0(\zeta) + \sum H_i^1 y'(\zeta_i) \tag{5.19}$$

其中 $a_i = \prod_{\substack{j=1\\ j\neq i}}^{N}\left(\frac{\zeta-\zeta_j}{\zeta_i-\zeta_j}\right)^2$，$b_i = \sum_{\substack{j=1\\ j\neq i}}^{N}\frac{1}{\zeta_i-\zeta_j}$.

对一维二点算例，$N=2, \zeta_1=0, \zeta_2=1$，

$$a_1 = \left(\frac{\zeta-\zeta_2}{\zeta_1-\zeta_2}\right)^2 = (\zeta-1)^2,\ b_1 = \frac{1}{\zeta_1-\zeta_2} = -1$$

$$a_2 = \left(\frac{\zeta-\zeta_1}{-1}\right)^2 = \zeta^2,\ b_2 = \frac{1}{\zeta_2-\zeta_1} = 1$$

所以

$$H_i^0 = a_i[1+2(\zeta_i-\zeta)b_i] \qquad H_i^1 = a_i(\zeta-\zeta_i)y'(\zeta_i)$$

$$H_1^0 = (\zeta-1)^2[1+2(-\zeta)(-1)] = (\zeta-1)^2(1+2\zeta) = 1-3\zeta^2+2\zeta^3$$

$$H_2^0 = 3\zeta^2-2\zeta^3$$

$$H_1^1 = \zeta-2\zeta^2+\zeta^3$$

$$H_2^1 = \zeta^3-\zeta^2$$

推广：一维元素两端要求二阶导数连续，H 必是 $2^2+1=5$ 次多项式，每一节点 $\bar{u}$ 的由三个幂级数组成。

$$\bar{u}=\sum_{i=1}^{2}(\phi_i^0\bar{u}_i+\phi_i^1\bar{u}'_i+\phi_i^2\bar{u}''_i)=\sum_{i=1}^{2}(H_i^0\bar{u}_i+H_i^1\bar{u}'_i+H_i^2\bar{u}''_i)$$

$$H_0^0=1-10\zeta^3+15\zeta^4+6\zeta^5$$

$$H_0^1=\zeta-6\zeta^3+8\zeta^4-3\zeta^5$$

$$H_0^2=\frac{1}{2}(\zeta^2-3\zeta^3+3\zeta^4-\zeta^5)$$

……

三、小结

1. 元素交界处（一维为节点，二维为边界线）

变量及其所要求的 $m=n-1$ 阶导数必须连续。这是协调性要求，它保证物理量的加权余量积分式唯一，边界上保持到几阶导数连续必须由方程的性质与解法确定，使用分部积分降阶可使求解简化。

2. 极限状态

当元素长度缩小到零时，变量 u 的极限可用 $\bar{u}$ 统一表示为 $\lim\limits_{\Delta\to 0}\|u-\bar{u}\|<\varepsilon$，其中 Δ 为元素长度。

5.1.4 迦辽金法的推广

一、问题的引入

我们从解具体工程实际问题着手，不从数学理论入手。

例 均匀材料棒的一维热传导问题。

$$k\frac{\mathrm{d}^2T}{\mathrm{d}x^2}=0 \tag{5.20}$$

边界条件： $x=0(i=1),\ T=T_0;\ x=L(i=N),\ T=T_L$

解析解：$T(x)=(T_L-T_0)\dfrac{x}{L}+T_0$。

有限元迦辽金法：$\overline{T}=\sum\limits_{i=1}^{N}\overline{T}_i\phi_i\approx\sum\limits_{k=1}^{N}\alpha_k\phi_k$

$$\int_{x_1}^{x_N}k\phi_j\frac{\mathrm{d}^2\overline{T}}{\mathrm{d}x^2}\mathrm{d}x=0,\ i,j=1,2,\cdots,n \tag{5.21}$$

其中 $k\dfrac{\mathrm{d}^2\overline{T}}{\mathrm{d}x^2}$ 是 ε，ϕ_j 是迦辽金法中的加权函数。如果解区分解成 N 个节点，则有 $N-2$ 个交界点。积分方程的最高阶次是“2”，因此要求交界点上的物理量 $\dfrac{\mathrm{d}\overline{T}}{\mathrm{d}x}$ 连续，必须采用较复杂的 Hermite 插值多项式。

分部积分：

$$k\phi_j(x)\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\bigg|_{x_1}^{x_N}-\int_{x_1}^{x_N}k\frac{\mathrm{d}\phi_i}{\mathrm{d}x}\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\mathrm{d}x=0$$

所以

$$\int_{x_1}^{x_N} k\frac{\mathrm{d}\phi_j}{\mathrm{d}x}\left\{\frac{\mathrm{d}\phi_i}{\mathrm{d}x}\right\}\mathrm{d}x\{\overline{T}_i\} = Q_N\phi_N - Q_1\phi_1$$

$$Q_1 = k\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\bigg|_{x=x_1},\ Q_N = k\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\bigg|_{x=x_N}$$

Q_1,Q_N表示始末端的热流，先将它当做是未定的。

在积分式中，$\{\overline{T}_i\} = (\overline{T}_1,\overline{T}_2,\cdots,\overline{T}_N)^T$，$\left\{\frac{\mathrm{d}\phi_i}{\mathrm{d}x}\right\} = \left(\frac{\mathrm{d}\phi_1}{\mathrm{d}x},\frac{\mathrm{d}\phi_2}{\mathrm{d}x},\cdots,\frac{\mathrm{d}\phi_N}{\mathrm{d}x}\right)^T$，$j$从1到$N$共$N$个方程。

上述方程组可写成：$\sum_{i=1}^{N} K_{ij}\overline{T}_i = \begin{cases} Q_1,\ j=1 \\ 0,\ j=2,3,\cdots\cdots \\ Q_N,\ j=N \end{cases}$

$$\{K\}\{\overline{T}\} = \{Q\},$$

$$\{Q\} = \{Q_1,0,\cdots,Q_N\}^T,$$

K是$N\times N$矩阵，$\overline{T}$是$N\times 1$矩阵。

$$K_{ij} = \int_{x_1}^{x_N} k\frac{\mathrm{d}\phi_j}{\mathrm{d}x}\frac{\mathrm{d}\phi_i}{\mathrm{d}x}\mathrm{d}x,\quad i,j=1,2,\cdots,N$$

N个代数方程解出N个$\overline{T}$，则方程有解。这种解称为弱解，因为使用分部积分使积分方程中最高阶导数降阶。

不使用Hermite插值公式，而用Lagrange插值函数为基函数，计算方便，但是$\overline{T}$的二阶导数不一定存在，而使原方程无法满足，从广义意义上称为弱解，求解时还必须引进边界值Q的具体信息，否则无法求解。

二、分部积分法求解的基本概念

(1) 自伴算子与自然边界条件、本质边界条件

考虑区域V的一组齐次方程$\overline{L}(u)=0$，$u=u(x)$，$x\in V$，$\overline{L}(u)$与另一个函数$w(x)$的内积，在求积分时进行分部积分，一直进行到积分号内对u的偏导数消失为止，得：

$$\langle\overline{L},w\rangle = \langle u,\overline{L}^*\rangle + \int_S [F^*(w)G(u) - F(u)G^*(w)]\mathrm{d}S \tag{5.22}$$

其中S是体积V的外表面，F、G是分部积分后自然形成的。

例如：$\overline{L}(u) = \alpha\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + \beta u$，$\quad 0<x<1$，$\alpha,\beta=\mathrm{const}$

$$\begin{aligned}\langle\overline{L}(u),w\rangle &= \int_0^1\left(\alpha\frac{\mathrm{d}^2u}{\mathrm{d}x^2}+\beta u\right)w\mathrm{d}x \\ &= \alpha\frac{\mathrm{d}u}{\mathrm{d}x}w\bigg|_0^1 + \int_0^1\beta uw\mathrm{d}x - \int_0^1\alpha\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}w}{\mathrm{d}x}\mathrm{d}x \\ &= \left[\alpha w\frac{\mathrm{d}u}{\mathrm{d}x} - u\frac{\mathrm{d}w}{\mathrm{d}x}\right]\bigg|_0^1 + \int_0^1 u\left(\alpha\frac{\mathrm{d}^2w}{\mathrm{d}x^2}+\beta w\right)\mathrm{d}x\end{aligned} \tag{5.23}$$

到此时u的导数在积分号内全部消失。与(5.22)式比较，有

$$\overline{L}^* = \overline{L} = \left(\alpha\frac{\mathrm{d}^2}{\mathrm{d}x^2}+\beta\right),\ F^*=\alpha,\ G=\frac{\mathrm{d}}{\mathrm{d}x},\ F=+1,\ G^*=\frac{\mathrm{d}}{\mathrm{d}x}$$

定义 凡分部积分，使得原微分号在积分式中消失得到$\overline{L}^*(w)$，称其是$\overline{L}$的伴随算子。如果$\overline{L}^*=\overline{L}$则称$\overline{L}$为自伴算子，此时也保证有$G=G^*$。可见，(5.23)中$\overline{L}$是自伴算

子。

分部积分使内积转置这种过程与矩阵的转置相对应，自伴算子的性质与矩阵的对称性质相通。

例 $A=[a_{ij}]$，若 $A^T=A$，则 $a_{ij}=a_{ji}$，A 为对称矩阵。对两个 n 维矢量 $\boldsymbol{w}$ 与 $\boldsymbol{u}$，若 $A=n\times n=A^T$，则

$$\langle \boldsymbol{w},\ A\boldsymbol{u}\rangle \equiv \langle \boldsymbol{u}, A^T\boldsymbol{w}\rangle$$

若边界响应为零，则有

$$\langle w,\ L(u)\rangle \equiv \langle u, \bar{L}^*(w)\rangle$$

可见自伴算子内积与对应矩阵的对称性相关。

结论：分部积分进行运算转置，除了判定算子是否自伴之外，还伴生不同类型的边界条件，规定 F 类边界条件 F^* 与 F 称为本质边界条件，它是边界上的变量值，属于第一类边界条件（即转换过程中它在边界上给定值）。G 类边界条件称为自然边界条件，与第二边值 Von Neumann 条件相对应，$G=\frac{\mathrm{d}}{\mathrm{d}x}\cdots$，通常求解时包含上面两类边界条件，可以证明自然边界条件在求解过程中自然满足。

(2) 非自伴情况经常存在，如 $\bar{L}(u)=\frac{\mathrm{d}^2u}{\mathrm{d}x^2}+\frac{\mathrm{d}u}{\mathrm{d}x}+u$。

$$\int_0^1 \bar{L}(u)w\mathrm{d}x=\int_0^1\left(\frac{\mathrm{d}^2w}{\mathrm{d}x^2}-\frac{\mathrm{d}w}{\mathrm{d}x}+w\right)\mathrm{d}x+\left[w\left(\frac{\mathrm{d}u}{\mathrm{d}x}+u\right)-u\frac{\mathrm{d}w}{\mathrm{d}x}\right]\Big|_0^1$$

所以 $\bar{L}\neq\bar{L}^*$，$G\neq G^*$。这就是非自伴算子，但无论如何随 $\bar{L}^*(w)$ 阶次的升高，自然边界条件与本质边界条件的阶次也会相应升高，但不会高于方程的最高阶次，流体力学大部分算子并非对称与正定的，是非自伴算子，因此计算更复杂些。

三、函数的连续程度与弱提法

前面讲 WRM 中的迦辽金法，主要用单一元素，对于边界点的连续性不必多说，但对一维多个元素或多维多元素，必须对边界点上的连续性进行深入分析。

设有微分方程

$$\bar{L}(u)=p,\ \text{边界条件}\ L(u)=g \tag{5.24}$$

(1) $u=u(x)$ 具有充分可微性且满足上述方程，$\bar{L}$ 若含有最高阶导数“n”，则 $u\in w_2^{n+1}$，w_2^n 是对函数的导数连续性加以限制的空间，称索柏洛夫空间。下标 2 表示平方可积的二次范数，“n”指有限导数的最高阶次。例如：

$$\| u \|_1=\int\left[u^2+\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2\right]\mathrm{d}x<\infty\Rightarrow\in w_2^1$$

$$\| u \|_2=\int\left[u^2+\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2+\left(\frac{\mathrm{d}^2u}{\mathrm{d}x^2}\right)^2\right]\mathrm{d}x<\infty\Rightarrow\in w_2^2$$

充分可微性即微分方程(5.24)中最高阶导数连续。

(2) $u(x)$ 局部可积，不是充分可微，即(5.24)中最高阶导数为 n，u 的 n 阶导数存在，不降阶直接求解即 $u\in w_2^n$，称为强解。

(3) 若 $\langle(\bar{L}(u)-p),w\rangle=0$，可用分部积分等方法求出 u 的数值分布，$u(x)$ 的连续性下降，对基函数（插值函数）在交界节点处要求下降，称为弱解，即存在导数间断的解。规

定凡是范数满足 $\|u\|_{\infty}=\int_V u^2\mathrm{d}V<\infty$ 的可积函数属于 L_2 范数空间的。

总之，微分方程 $a_1\dfrac{\mathrm{d}^n u}{\mathrm{d}x^n}+a_2\dfrac{\mathrm{d}^{n-1}u}{\mathrm{d}x^{n-1}}+\cdots=0$ 求经典解，u 的 $n+1$ 阶导数存在，n 阶连续，$u\in w_2^{n+1}$。

求强解：u 的 n 阶导数存在，$u\in w_2^n$。

求弱解：u 的 n 阶导数不存在，$u\in w_n^L$，$n=2,L=1$；$n=4,L=2$；$n=6,L=3$。

例如：

$$L(u)=\frac{\mathrm{d}^2u}{\mathrm{d}x^2}+u-x=0 \tag{5.25}$$

边界条件：$u=f,\quad x=0$；$u_x=g,\quad x=1$

迦辽金加权余量法求解：

$$\int\left(\frac{\mathrm{d}^2\bar{u}}{\mathrm{d}x^2}+u-x\right)\bar{w}\mathrm{d}x+\left(g-\frac{\mathrm{d}\bar{u}}{\mathrm{d}x}\right)\bar{w}|_{x=1}=0 \tag{5.26}$$

$\bar{u}\in w_2^2$，$\bar{u}$ 的一阶导数连续，二阶导数存在，$\bar{w}\in L_2$。

分部积分：

$$\int_0^1\left[(u-x)\bar{w}-\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}\bar{w}}{\mathrm{d}x}\right]\mathrm{d}x+g\bar{w}|_{x=1}=0 \tag{5.27}$$

$\bar{w}\in w_2^1$，$\bar{u}\in w_2^1$，只要求 0 阶导数连续，一阶导数存在。再分部积分：

$$\int\left[(u-x)\bar{w}-u\frac{\mathrm{d}^2\bar{w}}{\mathrm{d}x^2}\right]\mathrm{d}x-\left[g\bar{w}-u\frac{\mathrm{d}\bar{w}}{\mathrm{d}x}\right]\Bigg|_0^1=0,\ \bar{w}\in w_2^2,\ u\in L_2 \tag{5.28}$$

最常用的是(5.27)式，它可使对插值函数的要求降低；(5.28)式在迦辽金法中对插值函数的要求与(5.26)一致，没有多大好处。

在求解过程中，迦辽金法将基函数作为权函数。比较可见取(5.27)式只要求一阶连续，用 Lagrange 插值函数作为基函数正好，是合理的选择。

自然边界条件的满足与基函数的选取有关。

可见，在上例中，只要将 $x=0$ 处的 u 值给出，即可解出，即本质边界条件必须在求解中加入。自然边界条件是在分部积分求解中自然满足的边界条件。当边界条件不能自然满足时，那么用另一种加权函数求解边界条件的余量。

5.2 变分法(Variational)

5.2.1 变分法解微分方程的基本原理

一、泛函与变分

凡变量由一个或几个函数所确定称为泛函。

$$\begin{aligned}J_1[u(x)]&=\int_{x_0}^{x_1}F(u,u',x)\mathrm{d}x\\J_2[u(x)]&=\int_{x_0}^{x_1}[L(u)-p]\mathrm{d}x\end{aligned} \tag{5.29}$$

变分：研究泛函的极值问题，作法类同于对函数求导数。$\delta J[u(x)]=0$，但我们感兴趣的是 $\delta^2 J>0$，即极小值问题，这与矩阵的正定性相对应。

二、泛函 J 的 Euler 方程

1. 一维情况

$J=\int_{x_0}^{x_1}F(x,u,u',u'',\cdots,u^n)\mathrm{d}x$ 有极值存在。

$\delta J=\frac{\partial J}{\partial\alpha}\delta\alpha=0$，$\alpha$ 是 u 的某个增量形式，$\delta\alpha$ 值可以任取。极值存在的必要条件是：

$$\frac{\mathrm{d}F}{\mathrm{d}u}-\frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}F}{\mathrm{d}u'}+\frac{\mathrm{d}^2}{\mathrm{d}x^2}\frac{\mathrm{d}F}{\mathrm{d}u''}+\cdots+(-1)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}\frac{\mathrm{d}F}{\mathrm{d}u^n}=0 \tag{5.30}$$

这是一般形式的 Euler 方程。

2. 二维一阶情况

$$J=\iint F(x,y,u,u_x,u_y)\mathrm{d}x\mathrm{d}y=0 \tag{5.31}$$

对应的 Euler 方程是：

$$\frac{\partial F}{\partial u}-\frac{\partial}{\partial x}\frac{\partial F}{\partial u_x}-\frac{\partial}{\partial y}\frac{\partial F}{\partial u_y}=0 \tag{5.32}$$

二维高阶函数也可用上述方法延伸，得到 Euler 方程，但要注意交叉导数项的存在。

例 设

$$F(x,y,u,u_x,u_y)=\frac{1}{2}(u_x^2+u_y^2)-f(x,y)u \tag{5.33}$$

则 Euler 方程是：

$$-(u_{xx}+u_{yy})=f(x,y) \tag{5.34}$$

可见 Poisson 方程(5.34)是被积函数为(5.33)的泛函的 Euler 方程。

三、变分原理

现在关心的是逆问题即已知微分方程

$$\overline{L}(u)=p \tag{5.24}$$

$\overline{L}$ 是自伴线性正定算子，现在问若此方程是泛函取最小值的 Euler 方程其对应的泛函 J 是什么？

$$J(u)=\langle\overline{L}(u),u\rangle-2\langle p,u\rangle=\int_V[\overline{L}(u)u-2pu]\mathrm{d}V \tag{5.35}$$

即在 Hilbert 空间域中方程(5.24)有解 $u=u_1$ 存在，u_1 所满足的充分必要条件是(5.35)的 $J(u)$ 在 $u=u_1$ 处达到极小值。

证明 (1) 必要性

$\overline{L}(u_1)=p$，对 V 中任意的 $u=u_1+\eta$，则 $J(u)>J(u_1)$。

$$\begin{aligned}J(u)&=J(u_1+\eta)=\langle\overline{L}(u_1+\eta),u_1+\eta\rangle-2\langle p,u_1+\eta\rangle\\&=\langle\overline{L}(u_1),u_1\rangle+\langle\overline{L}(u_1),\eta\rangle+\langle\overline{L}(\eta),\eta\rangle+\langle\overline{L}(\eta),u_1\rangle-2\langle p,u_1\rangle-2\langle p,\eta\rangle\\&=J(u_1)+2\langle\overline{L}(u_1)-p,\eta\rangle+\langle\overline{L}(\eta),\eta\rangle\end{aligned}$$

所以 $J(u)>J(u_1)$，即 $J(u)$ 在 $u=u_1$ 处取极小值。

(2) 充分性

令 $u=u_1$ 时 $J(u)$ 极小，则 Euler 方程为 $\bar{L}(u_1)=p$，取 $\lim\limits_{\varepsilon\to 0}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}[J(u_1+\varepsilon\eta)]=0$，展开，由正定及对称性可证。

总之，相当一类微分方程 $\bar{L}(u_1)=p$，只要 $\bar{L}$ 是正定对称线性算子，可以转化为求解泛函的极小值这一变分问题，这样就为有限元求解开辟了一个新途径，即：

微分方程 → J → $\delta J=0$ → 求出 u 值。

一般解法：$J(u)=\int_V(\bar{L}(u)-2p)u\mathrm{d}V$

$$
\begin{aligned}
\delta J=\frac{\partial J}{\partial u}\delta u&=\int_V\{\delta u[\bar{L}(u)-2p]+u\bar{L}(\delta u)\}\mathrm{d}V\\
&=\langle\delta u,\bar{L}(u)\rangle-\langle\delta u,2p\rangle+\langle u,\bar{L}(\delta u)\rangle\\
&=2\langle\delta u,\bar{L}(u)\rangle-2\langle\delta u,p\rangle\\
&=2\langle\delta u,\bar{L}(u)-p\rangle=0
\end{aligned}
\tag{5.36}
$$

这就是迦辽金法。微分方程(5.24)有时可用变分法求解。当 $\bar{L}$ 是线性自伴正定算子时，此解法与迦辽金加权余量法是一致的。而迦辽金加权余量法则是变分法的推广，亦可用于求解非线性非正定算子的微分方程。

四、微分方程的变分求解

微分方程：

$$
\bar{L}(u)-p=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0,\qquad \boldsymbol{x}\in S \tag{5.37}
$$

边界条件：

$$
L(u)-q=\frac{\partial u}{\partial n}-q=0,\qquad \boldsymbol{x}\in l \tag{5.38}
$$

$\bar{L}$ 是线性、对称(自伴)、正定算子。

$$
\frac{\partial^2 u}{\partial x^2}\delta u=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\delta u\right)-\frac{\partial u}{\partial x}\delta\frac{\partial u}{\partial x}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\delta u\right)-\frac{\delta}{2}\left(\frac{\partial u}{\partial x}\right)^2
$$

$$
\frac{\partial^2 u}{\partial y^2}\delta u=\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\delta u\right)-\frac{\partial u}{\partial y}\delta\frac{\partial u}{\partial y}=\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\delta u\right)-\frac{\delta}{2}\left(\frac{\partial u}{\partial y}\right)^2
$$

$$
\begin{aligned}
\int\bar{L}(u)\delta u\mathrm{d}S&=\iint\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)\delta u\mathrm{d}x\mathrm{d}y\\
&=-\frac{\delta}{2}\int\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]\mathrm{d}x\mathrm{d}y+\int\left(\frac{\partial u}{\partial x}\delta u\mathrm{d}y+\frac{\partial u}{\partial y}\delta u\mathrm{d}x\right)\\
&=-\int\left(\frac{\partial u}{\partial x}\frac{\partial\delta u}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial\delta u}{\partial x}\right)\mathrm{d}x\mathrm{d}y+\int\left(\frac{\partial u}{\partial x}\delta u\frac{\mathrm{d}y}{\mathrm{d}l}+\frac{\partial u}{\partial y}\delta u\frac{\mathrm{d}x}{\mathrm{d}l}\right)\mathrm{d}l\\
&=-\int\left(\frac{\partial u}{\partial x}\frac{\partial\delta u}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial\delta u}{\partial x}\right)\mathrm{d}x\mathrm{d}y+\int\left(\frac{\partial u}{\partial x}\cos(n,x)+\frac{\partial u}{\partial y}\cos(n,y)\right)\delta u\mathrm{d}l\\
&=-\int\left(\frac{\partial u}{\partial x}\frac{\partial\delta u}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial\delta u}{\partial x}\right)\mathrm{d}x\mathrm{d}y+\int\left(\frac{\partial u}{\partial n}\right)\delta u\mathrm{d}l\\
&=\int L(u)\delta u\mathrm{d}l-\iint cD(u)D(\delta u)\mathrm{d}S
\end{aligned}
\tag{5.39}
$$

分部积分的一般形式，由(5.36)及(5.39)可得：

$$\delta J = 2\int(\bar{L}(u) - p)\delta u \mathrm{d}S = 0 \Rightarrow -\int[cD(u)D(\delta u) + p\delta u]\mathrm{d}x\mathrm{d}y + \int q\delta u \mathrm{d}l = 0 \tag{5.40}$$

(5.40)将自然边界条件包括进去了，但在求解中本质边界条件必须嵌入求解。在(5.24)中若 $p = 0$，则：

$$\int_S \bar{L}(u)\delta u \mathrm{d}S = \int_S L(u)\delta u \mathrm{d}l - \int_S cD(u)D(\delta u)\mathrm{d}S$$

对一般三维问题，如(5.24)，则有：

$$\delta J = -\int_V [cD(u)D(\delta u) + p\delta u]\mathrm{d}V + \int g\delta u \mathrm{d}S = 0 \tag{5.41}$$

这种降阶过程将自然边界条件包括进去了，可使边条处理简化。

5.2.2 Rayleigh-Ritz 法(瑞里一里兹法)

Rayleigh-Ritz 法用于解 $\delta J = 0$ 的二次泛函的极值问题。一方面，微分方程 $\bar{L}(u) = p$，当 $\bar{L}$ 是线性自伴算子(对称正定)，则将它视为 $J(u)$ 的 Euler 方程，由 $\delta J = 0$ 即可解出 u 来，且可证明与迦辽金法一致。

用变分方程求解灵活、方便，可利用数值分析的一些方法。

具体作法：

$$\begin{aligned} &\bar{L}(u) = p, \quad \boldsymbol{x} \in V \\ &L(u) = g, \quad \boldsymbol{x} \in S \\ &\bar{u} = \sum_{i=0}^{N} \alpha_i \phi_i \end{aligned} \tag{5.24}$$

ϕ_i 为基函数，必须满足本质边界条件，在节点处满足 $w_2^{(n)}$ 的要求，ϕ_i 也必须是线性无关的完备序列，如 $\bar{L}(u)$ 是线性二阶算子，降阶后是一阶，$D(\delta u) \in w_2^1$，$D(u) \in w_2^1(\bar{L}(u) \in w_2^2)$，此时降阶后可用 Lagrange 插值函数作基函数。由(5.40)得：

$$\delta J_j = -\int_V [cD(\bar{u})D(\delta\bar{u}_j) + p\delta\bar{u}_j]\mathrm{d}V + \int g\delta\bar{u}_j \mathrm{d}S = 0$$

$$\delta\bar{u}_j = \delta\alpha_i\phi_j$$

所以

$$\delta J_j = -\int[cD(\sum\alpha_j\phi_j)D(\delta\alpha_j\phi_j) + p\delta\alpha_j\phi_j]\mathrm{d}V + \int g\delta\alpha_j\phi_j \mathrm{d}S = 0$$

令 $\delta\alpha_j = 1$，则

$$\delta J_j = -\int[cD(\sum\alpha_j\phi_j)D(\phi_j) + p\phi_j]\mathrm{d}V + \int g\phi_j \mathrm{d}S = 0 \tag{5.42}$$

所以

$$\sum_{i=0}^{N}\int[cD(\sum\alpha_j\phi_j)D(\phi_j) + p\phi_j]\mathrm{d}V - \int g\phi_j \mathrm{d}S = 0,\ j = 0,1,\cdots,N$$

列出 $N + 1$ 个方程，解出 $N + 1$ 个 α_j。由于 ϕ_j 是已知基函数，于是 $\bar{u}$ 可解。

5.3 一维有限元法

下面开始就多个元素合成求解进行分析。



图 5.3

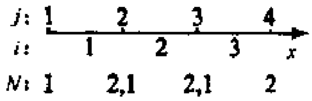


图 5.4

5.3.1 局部与总体有限元划分

定义域：$\Omega(0 \leqslant x \leqslant 1)$，边界：$S(x=0, x=1)$（见图5.3），将区域划分为若干子域，称为有限元。本例划分 3 个元，分别用 i 代表，每个元的交界点称节点。每个元素的模式为 $\bar{\Omega}_i = \Omega_i \cup S_i$，即一段区间加 2 个边界节点。

$$\begin{cases} S_{i=1} \cap S_{i=2} = S_{12} \\ \bar{\Omega} = \Omega \cup S = \sum_{i=1}^{3} \bar{\Omega}_i \end{cases}$$

图 5.4 将一个元素分成 3 个局部元素求解。三个标号表示法为：

元素单元号 i：$i = 1, 2, 3$；

总体节点号 j：$j = 1, 2, 3, 4$；

局部元素节点号 N：$N = 1, 2, \cdots, N$，在本例中 $N = 1, 2$。

为了表示第 i 个局部元素中的第 N 个节点，一般用 Z_N^i 表示。

若 Z_2^2 表示第二个局部元素的第二个节点 $\Rightarrow j = 3$（表示总体节点为 3）。其表达式为：

$$Z_N^i = \Delta_{Nj}^i Z_j \tag{5.43}$$

Z_j 表示总体节点数号，Z_N^i 表示局部单元数节点号。Δ_{Nj}^i 称为布尔矩阵，定义为

$$\Delta_{Nj}^i = \begin{cases} 1 & N \Rightarrow j \\ 0 & N \nRightarrow j \end{cases} \tag{5.44}$$

$N \Rightarrow j$ 表示局部元素的当地节点号与相应的总体节点号对应，例如 $\Delta_{1,3}^3 = 1$，$\Delta_{2,3}^3 = 0$。

5.3.2 局部元素与总体元素的基函数

使用迦辽金法要选择好基函数，如前所述，对一个有多个元素组成的总体有限元，按微分方程的特点，用 Lagrange 或 Hermite 插值函数比较合理。我们的思想是先在某一个局部元素上建立一般性的求解公式，再利用总体合成加以组合。

一、线性分布元素

由图 5.1，如取两个节点，$N=1$ 相当于(5.15)中的 ϕ_0，$N=2$ 相当于(5.15)中的 ϕ_1：

$$\phi_{N=1} = \frac{1-\zeta}{2}, \quad \phi_{N=2} = \frac{1+\zeta}{2} \tag{5.15a}$$

原点在 $\zeta=0$(中点处)，每一个局部元素上的试验函数为：

$$\bar{u}^i=\sum_{N=1}^{2}\phi_N^i\bar{u}_N^i=(\phi_1^i,\phi_2^i)\begin{pmatrix}\bar{u}_1^i\\ \bar{u}_2^i\end{pmatrix}$$

$\bar{u}_N^i$ 是第 i 个局部元素节点 N 上的近似解(待求)，即当地值，它与总体节点值的关系为 $\bar{u}_N^i=\Delta_{Nj}^i\bar{u}_j$，所以

$$\bar{u}=\sum_{i=1}^{3}\bar{u}^i=\sum_{i=1}^{3}\phi_N^i\Delta_{Nj}^i\bar{u}_j=\phi_j\bar{u}_j \tag{5.45}$$

这里 j 应求和。总体基函数为：

$$\phi_j=\sum_{i=1}^{3}\phi_N^i\Delta_{Nj}^i,\quad 例如\ \phi_3=\phi_2^2+\phi_1^3 \tag{5.45a}$$

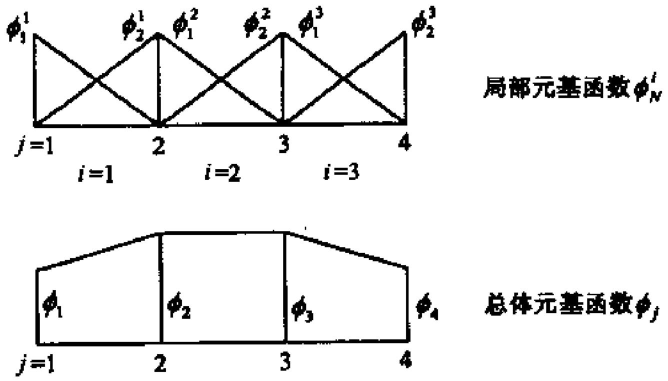


图 5.5

二、二次插值函数

由(5.16)式得：

$$\phi_1^i=\frac{\zeta(1-\zeta)}{2},\quad \phi_2^i=1-\zeta^2,\quad \phi_3^i=\frac{\zeta(1+\zeta)}{2} \tag{5.46}$$

一个局部元素 3 个节点，其它布尔矩阵及及总体与局部元素的公式为(5.45)与(5.45a)式，但 Δ_{Nj}^i 的确定要与总体节点号和局部节点号相匹配。

5.3.3 迦辽金(Galerkin)加权余量法解有限元问题

思路：先按前面已经讲过的办法写出局部元的求解式，然后按照(5.45)与(5.45a)式建立总体元与局部元的计算关系式，则总体元可解。

微分方程 $\qquad \bar{L}(u)-p=0 \qquad (5.24)$

边界条件 $\qquad L(u)-g=0$

一、总体有限元的解法

令：

$$\bar{u}=\phi_j\bar{u}_j=(\phi_1\quad\phi_2\quad\cdots)\begin{pmatrix}\bar{u}_1\\ \bar{u}_2\\ \vdots\end{pmatrix},\quad \phi_j=\sum_{i=1}^{I}\phi_N^i\Delta_{Nj}^i \tag{5.47}$$

则：

$$\int_{\Omega}(\overline{L}(\bar{u})-p)\phi_k\mathrm{d}\Omega-\int_{S}(L(\bar{u})-g)\phi_k\mathrm{d}s=0 \tag{5.48}$$

降阶与分部积分：

$$\int_{V}cD(\bar{u})D(\phi_k)\mathrm{d}\Omega+\int_{\Omega}p\phi_k\mathrm{d}\Omega-\int g\phi_k\mathrm{d}S=0 \tag{5.49}$$

若分析的问题是第一边界条件问题，g 项不存在，但在求解时本质边界条件作为强制约束条件嵌入求解；若是混合边界条件，自然边界条件在降阶的分部积分时得到的公式会自动得到满足，也只须加上第一边值的限制即可。必须指出，有时求解二、三维问题时插值函数并不包括边界上的情况，边界上插值函数的要求并不一致，要另外给出。一维问题二者是一致的。

最后将 $\bar{u}=\phi_j\bar{u}_j$ 代入(5.49)式得

$$\int_{\Omega}cD(\phi_j)D(\phi_k)\mathrm{d}\Omega\{\bar{u}_j\}=-\int_{\Omega}p\phi_k\mathrm{d}\Omega+\int g\phi_k\mathrm{d}S \tag{5.50}$$

写出矩阵形式，若 $g=0$，得

$$(A_{jk})(\bar{u}_j)=F_k \tag{5.51}$$

$$F_k=-\int p\phi_k\mathrm{d}\Omega \tag{5.52}$$

这里$(A_{jk})=\int_{\Omega}cD(\phi_j)D(\phi_k)\mathrm{d}\Omega$ 为 $J\times J$ 矩阵，它是 $\overline{L}(\bar{u})$ 分部积分后得到的系数矩阵，称为总体系数矩阵，对全域(i 个元素)都成立；F_k 是 $J\times 1$ 的列矢量，称为总体输入矢量，$j,k=1,2,\cdots,J$，于是可以解出 $\bar{u}_j(j=1,2,\cdots,J)$。

这个总体有限元不必另外进行计算，只须将与此微分方程一致的局部有限元的当地元素方程列出，然后根据总体元与局部元的匹配关系式即可求解。

二、局部有限元方程

微分方程与边界条件是(5.24)，有限元方程是前面讲过的(5.48)式。下面举例分析。

令 $\overline{L}(u)-p=\dfrac{\mathrm{d}^2u}{\mathrm{d}x^2}-a^2u-f(x)=0$，边界条件是本质边界条件。

按(5.48)式：$\int_0^h\left[\dfrac{\mathrm{d}^2\bar{u}}{\mathrm{d}x^2}-a^2\bar{u}-f(x)\right]\phi_M^i\mathrm{d}x=0$，其中 h 是局部元长度。

其中 $\bar{u}=\sum\phi_N^i\bar{u}_N^i$.

分部积分有限元方程得：

$$\int_0^h\left(\frac{\mathrm{d}\phi_N^i}{\mathrm{d}x}\frac{\mathrm{d}\phi_M^i}{\mathrm{d}x}+a^2\phi_N^i\phi_M^i\right)\mathrm{d}x\{u_N^i\}=-\int_0^h f(x)\phi_M^i\mathrm{d}x$$

写成：

$$A_{NM}^i\bar{u}_N^i=F_M^i \tag{5.53}$$

其中

$$A_{NM}^i=\int_0^h\left(\frac{\mathrm{d}\phi_N^i}{\mathrm{d}x}\frac{\mathrm{d}\phi_M^i}{\mathrm{d}x}+a^2\phi_N^i\phi_M^i\right)\mathrm{d}x$$

$$F_M^i=-\int_0^h f(x)\phi_M^i\mathrm{d}x$$

对于用线性 Lagrange 插值函数为基函数有：

$$\phi_1^i = \frac{(1-\zeta)}{2}, \quad \phi_2^i = \frac{(1+\zeta)}{2}, \quad \zeta = 2\bar{x} - 1, \ \bar{x} = \frac{x}{h}$$

所以局部系数矩阵 A_{NM}^i 为：

$$A_{NM}^i = \frac{1}{h}\int_{-1}^{1}\left(\frac{\mathrm{d}\phi_N^i}{\mathrm{d}\zeta}\frac{\mathrm{d}\phi_M^i}{\mathrm{d}\zeta} + \frac{h}{2}a^2\phi_N^i\phi_M^i\right)\mathrm{d}\zeta, \quad N,M = 1,2$$

$$A_{NM}^i = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{h} + \frac{ha^2}{3} & \frac{ha^2}{6} - \frac{1}{h} \\ \frac{ha^2}{6} - \frac{1}{h} & \frac{1}{h} + \frac{ha^2}{3} \end{pmatrix} \tag{5.54}$$

令 $f(x) = H$ 为常数，则：

$$F_M^i = -\int_{-1}^{1}\phi_M H\frac{h}{2}\mathrm{d}\zeta = (-H)\begin{pmatrix} \frac{h}{2} \\ \frac{h}{2} \end{pmatrix} \tag{5.55}$$

称为局部输入矢量。

本节整个计算在前面 Galerkin 法中已全部学会，其它形式的微分方程完全可用类似于上述的方法。

三、总体有限元与局部有限元的关系

$$(A_{jk})(\bar{u}_j) = F_k, \qquad j,k = 1,2,\cdots \tag{5.51}$$

对上面的例子有：

$$A_{jk} = \int_0^l\left(\frac{\mathrm{d}\phi_j}{\mathrm{d}x}\frac{\mathrm{d}\phi_k}{\mathrm{d}x} + a^2\phi_j\phi_k\right)\mathrm{d}x, \qquad j = 1,2,3,4; k = 1,2,3,4 \tag{5.56}$$

其中由(5.45a)式得：

$$\phi_j = \sum_{i=1}^{I}\phi_N^i\Delta_{Nj}^i, \quad \phi_k = \sum_{i=1}^{I}\phi_M^i\Delta_{Mk}^i \tag{5.57}$$

代入(5.56)得：

$$A_{jk} = \sum_{i=1}^{I}\int_0^h\left(\frac{\mathrm{d}\phi_N^i}{\mathrm{d}x}\frac{\mathrm{d}\phi_M^i}{\mathrm{d}x} + a^2\phi_N^i\phi_M^i\right)\mathrm{d}x(\Delta_{Nj}^i\Delta_{Mj}^i) = \sum_{i=1}^{I}A_{NM}^i\Delta_{Nj}^i\Delta_{Mj}^i \tag{5.58}$$

即：

$$a_{jk} = \sum_{i=1}^{I}a_{NM}^i\Delta_{Nj}^i\Delta_{Mj}^i,$$

$$F_k = \sum_{i=1}^{I}F_M\Delta_{Mk}^i \tag{5.59}$$

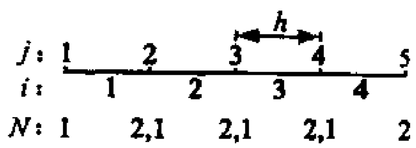


图 5.6

a_{jk}, a^i_{NM} 分别为 A_{jk}, A^i_{NM} 中各元素。应用以前的关于布尔矩阵的定义式，可以写出 A_{jk} 的表达式：

$$A_{jk}=\begin{bmatrix} a^1_{11} & a^1_{12} & & \\ a^1_{21} & a^1_{22}+a^2_{11} & a^2_{12} & \\ & a^2_{21} & a^2_{22}+a^3_{11} & a^3_{12} \\ & & a^3_{21} & a^3_{22} \end{bmatrix}$$

$$=\begin{bmatrix} \frac{1}{h}+\frac{ha^2}{3} & \frac{a^2h}{6}-\frac{1}{h} & & \\ \frac{ha^2}{6}-\frac{1}{h} & \frac{2}{h}+\frac{2}{3}ha^2 & \frac{a^2h}{6}-\frac{1}{h} & \\ & \frac{ha^2}{6}-\frac{1}{h} & \frac{2}{h}+\frac{2ha^2}{3} & \frac{a^2h}{6}-\frac{1}{h} \\ & & \frac{ha^2}{6}-\frac{1}{h} & \frac{1}{h}+\frac{ha^2}{3} \end{bmatrix}$$

$$F_k=\sum_{i=1}^{I}F^i_M\Delta^i_{Mk}=\begin{Bmatrix} F^1_1 \\ F^1_2+F^2_1 \\ F^2_2+F^3_1 \\ F^3_2 \end{Bmatrix}=(-H)\begin{Bmatrix} h/2 \\ h \\ h \\ h/2 \end{Bmatrix}$$

给出定解条件，就可求解。

例 令 $a^2=1$，$f=0$，$h=0.5$，$u(0)=1$，$u(1)=7.329$，则：

$$A^i_{NM}=\begin{bmatrix} 2.1667 & -1.9167 \\ -1.9167 & 2.1667 \end{bmatrix}$$

$$A_{jk}=\begin{bmatrix} 2.1667 & -1.9167 & \\ -1.9167 & 4.3334 & -1.9167 \\ & \ddots & \ddots \\ & & \end{bmatrix}$$

因为解域是 $0\leqslant x\leqslant 2$，于是分成 4 个元，每个元的长度 $h=0.5$，有 5 个总体节点见图 5.6，所以：

$$\begin{bmatrix} 2.1667 & -1.9167 & \\ -1.9167 & 4.3334 & -1.9167 \\ & \ddots & \ddots \\ & & \end{bmatrix}\begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \bar{u}_3 \\ \bar{u}_4 \\ \bar{u}_5 \end{bmatrix}=0$$

必须将本质边界条件强制插入，得：

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
-1.9167 & 4.3334 & -1.9167 & 0 & 0 \\
0 & -1.9167 & 4.3334 & -1.9167 & 0 \\
0 & 0 & -1.9167 & 4.3334 & -1.9167 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\bar{u}_1 \\ \bar{u}_2 \\ \bar{u}_3 \\ \bar{u}_4 \\ \bar{u}_5
\end{bmatrix}
=
\begin{bmatrix}
1 \\ 0 \\ 0 \\ 0 \\ 7.329
\end{bmatrix}
$$

解得：$\bar{u}_2 = 1.637$，$\bar{u}_3 = 2.692$，$\bar{u}_4 = 4.456$。

四、小结

(1) 一维问题用局部元素的 A^i_{NM} 在各节点上由(5.58)和(5.59)式分别求得总体系数矩阵和总体输入变量；

(2) 总体节点增加只将矩阵延伸，使矩阵阶数增大，通常带状对角元素值不变；

(3) 基函数将大大影响计算工作量；

(4) 边界条件约束必须在求解过程中仔细分析，自然边界条件在分部积分时要加上，其影响自然满足；本质边值必须在求解中强制嵌入，不然解无意义。

5.4 二维有限元法

从 Galerkin 法的基本原则与推导过程可知，一维与二维完全一样，只是积分域从线变成了面。求解方法是用局部元素与总体元相匹配的方法。现在的问题是：

- 二维问题局部元素如何选择？
- 选择怎样的基函数？
- 边界条件如何处理？

首先将前面求解的最基本的公式罗列如下：

$$(A_{jk})(\bar{u}_j) = F_k \tag{5.51}$$

$$A_{jk} = \sum_{i=1}^{I} A^i_{NM}\Delta^i_{Nj}\Delta^i_{Mj} \tag{5.58}$$

$$F_k = \sum_{i=1}^{I} F^i_M\Delta^i_{Mk} \tag{5.59}$$

5.4.1 矩形元问题的求解

一、有限元的划分

二维问题将解域划分为许多元素，元素间的边界由节点相联系，节点上的未知量是要求解的值，每个局部元素上基函数必须满足相容性、协调性的要求，即由微分方程最高阶导数决定节点处变量连续性的要求，最常用的元素形状是三角形与矩形。

方程的近似解是对元素来建立的，每个元素内通过插值函数 ϕ 来求解元素节点处的 $\bar{u}_j$ 值及其导数 $\bar{u}'_j$ 值等。

二、二维矩形局部元中的基函数

如用 Lagrange 线性插值函数作为基函数，保证节点处连续性的要求，只须将一维 Lagrange插值式匹配，即可得矩阵元基函数。

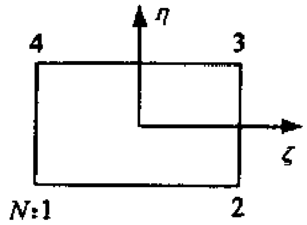


图 5.7

图 5.7 中注上了局部元节点，即 $N=1,2,3,4$。可利用(5.15)在 ζ,η 方向的匹配得：

$$\phi_1=\frac{1-\zeta}{2}\frac{1-\eta}{2},\quad \phi_2=\frac{1+\zeta}{2}\frac{1-\eta}{2},$$

$$\phi_3=\frac{1+\zeta}{2}\frac{1+\eta}{2},\quad \phi_4=\frac{1-\zeta}{2}\frac{1+\eta}{2}$$

写成统一表达式：

$$\phi_N=\frac{1+\zeta\zeta_N}{2}\frac{1+\eta\eta_N}{2}$$

其中 ζ_N,η_N 分别是局部元节点 N 处 ζ,η 的当地值，如 $N=1$ 处 $\zeta=-1,\eta=-1$。

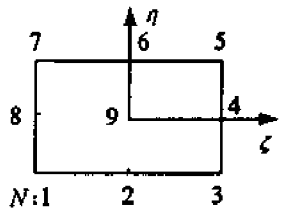


图 5.8

二次插值：其在某一方向中有三个节点，加一个中心节点则由图 5.8 所示的 9 个节点。此时所得的基函数分别是 ζ,η 的二次函数，函数的确定由与(5.46)的匹配而得：

$$\phi_N=\frac{\zeta\zeta_N}{4}\eta\eta_N(1+\eta\eta_N)(1+\zeta\zeta_N),\ N=1,3,5,7$$

$$\phi_N=\frac{1}{2}\eta\eta_N(1+\eta\eta_N)(1-\zeta^2),\ N=2,6$$

$$\phi_N=\frac{1}{2}\zeta\zeta_N(1+\zeta\zeta_N)(1-\eta^2),\ N=4,8$$

$$\phi_N=(1+\zeta^2)(1-\eta^2),\ N=9$$

如用 Hermite 多项式插值作为基函数也可类似于 Lagrange 基函数的匹配法用之。

三、局部元素 Poisson 方程求解

该方法是指 Poisson 方程用迦辽金加权余量法求解时 A^i_{NM},F_M 的求解。

方程：

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=Q_N=\text{const} \tag{5.60}$$

边界条件：

$$u(x,1) = u(0,y) = 1$$
$$u_y(x,0) = u_x(1,y) = 1$$

其图形见图 5.7,解区为 $0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1$。

1. 试验函数

$\bar{u}(x,y) = \bar{u}_N \phi_N$ 从 $N = 1$ 到 4 求和。

$$\phi_N = \frac{1+\zeta\zeta_N}{2}\frac{1+\eta\eta_N}{2}, \quad \zeta = 2\bar{x} - 1, \quad \eta = 2\bar{y} - 1$$

2. 迦辽金加权余量法的分部积分

$$\int_\Omega (\bar{u}_{xx} + \bar{u}_{yy} - Q)\phi_M \mathrm{d}\Omega - \int_S \left(\frac{\partial \bar{u}}{\partial n} - g\right)\mathrm{d}l = 0$$

$$\int_\Omega \left(\frac{\partial \phi_N}{\partial x}\frac{\partial \phi_M}{\partial x} + \frac{\partial \phi_N}{\partial y}\frac{\partial \phi_M}{\partial y}\right)\mathrm{d}x\mathrm{d}y\{\bar{u}_N^i\} = \int_l g\mathrm{d}l - \iint_\Omega Q\phi_M \mathrm{d}x\mathrm{d}y$$

$$\int_{-1}^{1}\int_{-1}^{1}\left(\frac{\partial \phi_N}{\partial \zeta}\frac{\partial \phi_M}{\partial \zeta} + \frac{\partial \phi_N}{\partial \eta}\frac{\partial \phi_M}{\partial \eta}\right)\mathrm{d}\zeta\mathrm{d}\eta\{\bar{u}_N^i\} = \int_{-1}^{1}\int_{-1}^{-1}\frac{Q}{4}\phi_M \mathrm{d}\zeta\mathrm{d}\eta$$

3. 写出 A_{NM}^i, F_M^i

$$A_{NM}^i = \begin{bmatrix} a & -b & -c & -b \\ -b & a & -b & -c \\ -c & -b & a & -b \\ -b & -c & -b & a \end{bmatrix} \tag{5.61}$$

这里 $a = \frac{2}{3}, b = \frac{1}{6}, C = \frac{1}{3}$。

$$F_M^i = (-Q_N)\begin{Bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{Bmatrix} \tag{5.62}$$

4. 嵌入边界条件

如果要求解,必须将边界条件加入。但由于这是局部元素的求解,通常不必将边界条件嵌入,只作为总体元求解的基础,在总体元求解时边界条件的处理必不可少。

四、总体有限元求解

Poisson 方程:

$$\bar{L}(u) - p = \Delta u - Q = 0 \tag{5.60}$$

边界条件:

$$u(x,2) = 1, \ u(0,y) = 1, \ u_x|_{x=2} = 0, \ u_y|_{y=0} = 0$$

其中

$$Q = Q_w \delta(x-1)\delta(y-1) \tag{5.63}$$

δ 是 Dirac 函数。

1. 作图并划分元素,注明标号

由边界条件 $\bar{u}_3 = \bar{u}_6 = \bar{u}_9 = \bar{u}_1 = \bar{u}_2 = 1$,只要求出 $\bar{u}_4, \bar{u}_5, \bar{u}_7, \bar{u}_8$ 即可。

基本公式为(5.51),(5.58)和(5.59)。

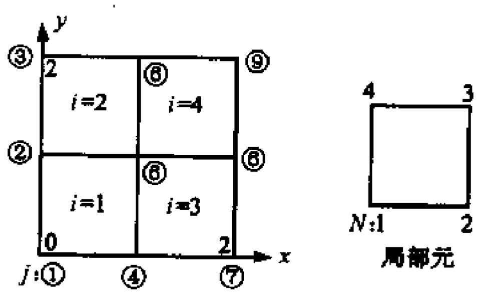


图 5.9

其中

$$A'_{NM}=\begin{bmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\\a_{41}&a_{42}&a_{43}&a_{44}\end{bmatrix}=\begin{bmatrix}\frac{2}{3}&-\frac{1}{6}&-\frac{1}{3}&-\frac{1}{6}\\-\frac{1}{6}&\frac{2}{3}&-\frac{1}{6}&-\frac{1}{3}\\-\frac{1}{3}&-\frac{1}{3}&\frac{2}{3}&-\frac{1}{6}\\-\frac{1}{6}&-\frac{1}{3}&-\frac{1}{6}&\frac{2}{3}\end{bmatrix}$$

这是已知局部元的结果，即(5.61)式，

$$F'_M=Q_w\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

即(5.62)式。

2. 求 $A_{jk}(9\times 9)$

根据(5.58)式(其中 $I=4$)，最后求得：

$$A_{jk}=\begin{bmatrix}2/3&-1/6&0&-1/6&-1/3&0&0&0&0\\&4/3&-1/6&-1/3&-1/3&-1/3&0&0&0\\&&2/3&0&-1/3&-1/6&0&0&0\\\text{对}&&&4/3&-1/3&0&-1/6&-1/3&0\\&&&&8/3&-1/3&-1/3&-1/3&-1/3\\&&&&&4/3&0&-1/3&0\\&&&&&&2/3&-1/6&0\\&&&&&&&4/3&-1/6\\&&&&\text{称}&&&&2/3\end{bmatrix}$$

$$F_k=-\iint Q_w\delta(x-1)\delta(y-1)\phi_k\mathrm{d}\Omega=-Q_w\phi_k|_{x=1,y=1}=-Q_w\phi_5$$

3. 加入边界条件

加入本质边界条件后由(5.51)可得：

$$\begin{bmatrix} 4/3 & -1/3 & -1/6 & -1/3 \\ & 8/3 & -1/3 & -1/3 \\ \text{对} & & 2/3 & -1/6 \\ & \text{称} & & 4/3 \end{bmatrix} \begin{bmatrix} \bar{u}_4 \\ \bar{u}_5 \\ \bar{u}_7 \\ \bar{u}_8 \end{bmatrix} = \begin{bmatrix} 1/2 \\ Q_w + 5/3 \\ 0 \\ 1/2 \end{bmatrix}, \quad \text{其中 } Q_w = 1$$

解得 $(\bar{u}_4 \quad \bar{u}_5 \quad \bar{u}_7 \quad \bar{u}_8)^T = (0.783 \quad 0.525 \quad 0.655 \quad 0.783)^T$.

而根据解析解可得：$(u_4 \quad u_5 \quad u_7 \quad u_8)^T = (0.756 \quad 0.127 \quad 0.719 \quad 0.756)^T$，比较可知，$u_5$ 误差较大，因为对复杂的温度分布，用线性分布 Lagrange 插值函数作为基函数不太合适，用二次插值函数作为基函数可使解更合理，但计算复杂。

上面已将求解过程与方法说清楚了。

5.4.2 三角元总体有限元求解

一、基函数的定义

若解域是一个非正规的几何外形，则用三角形元素比较合适。

这里介绍在节点上 $u \in w_2^{(1)}$ 的情况，即用线性插值函数的情况，$u \in w_2^{(2)}$ 的 Hermite 族基函数解法相同，但复杂得多，就不再介绍了。在一维与二维矩形元中，用 x,y，即空间坐标的参数作为基函数的自变量 $\phi = f(x,y)$，但非正规形的坐标不易确定，通常用自然坐标系，即面积坐标系来确定基函数。

对于三角形元：

$$\bar{u} = \sum \phi_N \bar{u}_N \tag{5.64}$$

$\phi_N = L_N = \dfrac{S_N}{S}$，$S$ 是三角形的总面积。

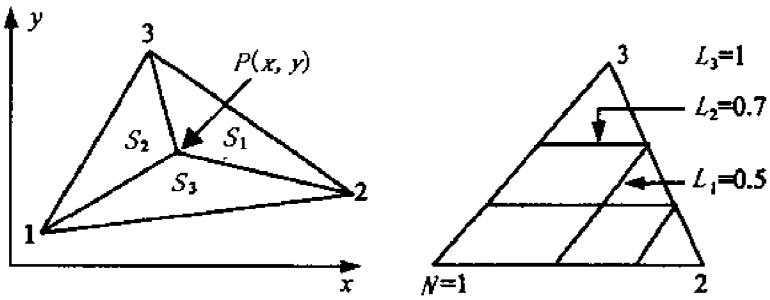


图 5.10

现在将三角形中任意点 $P(x,y)$ 作顶点连线，以顶点对应边为底边的三角形冠以顶点的标号，得面积 S_1,S_2,S_3，无因次面积：$L_N = \dfrac{S_N}{S}$。用 L_N 作为基函数的变量，所以 $\phi_N = L_N$，显然，在顶点处，$L_N = 1$，而在三角形的其它两个端点为 0，符合 Lagrange 插值函数的要求。且 $L_1+L_2+L_3=1$。

$$\bar{u} = \sum_{N=1}^{3} L_N \bar{u}_N \tag{5.65}$$

当三角形元的三个顶点由 $P_1(x_1,y_1)$，$P_2(x_2,y_2)$，$P_3(x_3,y_3)$ 确定，由解析几何得整个

三角形的总面积为：

$$S = \frac{1}{2}\det\begin{bmatrix}1 & x_1 & y_1\\ 1 & x_2 & y_2\\ 1 & x_3 & y_3\end{bmatrix} \tag{5.66}$$

局部面积为：

$$S_2 = \frac{1}{2}\det\begin{bmatrix}1 & x_1 & y_1\\ 1 & x & y\\ 1 & x_3 & y_3\end{bmatrix} \tag{5.67}$$

S_1,S_3 由此类推(x,y 为三角形中 $P(x,y)$ 点位置)。

所以 $L_2 = \frac{S_2}{S}$,同理可求得 L_1,L_3，最后得基函数：

$$\begin{bmatrix}L_1\\ L_2\\ L_3\end{bmatrix} = \frac{1}{2S}\begin{bmatrix}y_3x_2 - y_2x_3 & y_2 - y_3 & x_3 - x_2\\ y_1x_3 - y_3x_1 & y_3 - y_1 & x_1 - x_3\\ y_2x_1 - y_1x_2 & y_1 - y_2 & x_2 - x_1\end{bmatrix}\begin{bmatrix}1\\ x\\ y\end{bmatrix} \tag{5.68}$$

二、Poisson 方程的局部元素的矩阵

若 $\overline{L}(u) - p = \Delta u - Q = 0$，迦辽金 WRM 法分部积分得：

$$\iint\left(\frac{\partial\phi_N}{\partial x}\frac{\partial\phi_M}{\partial x} + \frac{\partial\phi_N}{\partial y}\frac{\partial\phi_M}{\partial y}\right)\mathrm{d}x\mathrm{d}y\{\tilde{u}_N\} = -\iint Q\phi_M\mathrm{d}x\mathrm{d}y$$

$$\iint\left(\frac{\partial L_N}{\partial x}\frac{\partial L_M}{\partial x} + \frac{\partial L_N}{\partial y}\frac{\partial L_M}{\partial y}\right)\mathrm{d}\Omega\{\tilde{u}_N^i\} = A_{NM}^i\{\tilde{u}_N^i\} = F_M^i = -\iint Q\phi_M\mathrm{d}x\mathrm{d}y \tag{5.69}$$

因为 $L_1 + L_2 + L_3 = 1$,所以 $L_3 = L_3(L_1,L_2) = 1 - L_1 - L_2$，所以

$$\frac{\partial L_3}{\partial x} = -\left(\frac{\partial L_2}{\partial x} + \frac{\partial L_1}{\partial x}\right),\quad \frac{\partial L_3}{\partial y} = -\left(\frac{\partial L_2}{\partial y} + \frac{\partial L_1}{\partial y}\right)$$

$$I = I_x + I_y = \iint\frac{\partial L_N}{\partial x}\frac{\partial L_M}{\partial x}\mathrm{d}x\mathrm{d}y + \iint\frac{\partial L_N}{\partial y}\frac{\partial L_M}{\partial y}\mathrm{d}x\mathrm{d}y = A_{NM}^i \tag{5.70}$$

先分析 I_x：

$$\frac{\partial L_N}{\partial x}\frac{\partial L_M}{\partial x} = \left\{\frac{\partial L_N}{\partial L_1}\frac{\partial L_1}{\partial x} + \frac{\partial L_N}{\partial L_2}\frac{\partial L_2}{\partial x}\right\}\cdot\left\{\frac{\partial L_M}{\partial L_1}\frac{\partial L_1}{\partial x} + \frac{\partial L_M}{\partial L_2}\frac{\partial L_2}{\partial x}\right\}$$

$$= \left[\frac{\partial L_N}{\partial L_1}(y_2 - y_1) + \frac{\partial L_N}{\partial L_2}(y_3 - y_1)\right]\cdot\left[\frac{\partial L_M}{\partial L_1}(y_2 - y_3) + \frac{\partial L_M}{\partial L_2}(y_3 - y_1)\right]$$

最后通过一番归纳,可得

$$\beta = \frac{\partial L_N}{\partial x}\frac{\partial L_M}{\partial x} = \frac{1}{4S^2}(y_{N+1} - y_{N+2})(y_{M+1} - y_{M+2}) \tag{5.71}$$

这里 N 为循环量,N 从 0,1,2,3 再到 1。M 类似。所以

$$I_x = \iint_\Omega\frac{\partial L_N}{\partial x}\frac{\partial L_M}{\partial x}\mathrm{d}\Omega = \beta\iint_\Omega\mathrm{d}\Omega = \beta S = \frac{1}{4S}(y_{N+1} - y_{N+2})(y_{M+1} - y_{M+2}) \tag{5.72}$$

这是由于 β 是常数，面对整个区域积分最后结果显然是区域的面积。同理可得：

$$I_y = \iint_\Omega\frac{\partial L_N}{\partial y}\frac{\partial L_M}{\partial y}\mathrm{d}\Omega = \frac{1}{4S}(x_{N+1} - x_{N+2})(x_{M+1} - x_{M+2}) \tag{5.73}$$

最后得

$$A'_{NM} = \frac{1}{4S}\begin{Bmatrix} b_1^2 & b_1b_2 & b_1b_3 \\ 对 & b_2^2 & b_2b_3 \\ & 称 & b_3^2 \end{Bmatrix} + \frac{1}{4S}\begin{Bmatrix} a_1^2 & a_1a_2 & a_1a_3 \\ 对 & a_2^2 & a_2a_3 \\ & 称 & a_3^2 \end{Bmatrix} \tag{5.74}$$

其中：$a_l = x_n - x_m$，$b_l = y_n - y_m$。l,m,n 的对应关系为 $l = 1,2,3$；$m = 2,3,1$；$n = 3,1,2$。

分别列出三个方程可解三个未知量 $\bar{u}'_1,\bar{u}'_2,\bar{u}'_3$。求解过程如下：

局部输入矢量 $F'_M = -\iint_{\Omega} Q\phi'_M \mathrm{d}\Omega = -\iint_{\Omega} QL_M \mathrm{d}\Omega$，

张捷迁证明了 $\iint_{\Omega} L_1^{m1}L_2^{m2}L_3^{m3}\mathrm{d}\Omega = \frac{m1!\cdot m2!\cdot m3!}{(m1+m2+m3+2)!}2S$

在这里，$m1 = 1, m2 = m3 = 0$，所以：

$$F'_M = -Q\frac{1}{3!}2S = -\frac{QS}{3} = -\frac{QS}{3}\begin{Bmatrix}1\\1\\1\end{Bmatrix} \tag{5.74a}$$

$$\{A'_{NM}\}\begin{Bmatrix}\bar{u}_1\\ \bar{u}_2\\ \bar{u}_3\end{Bmatrix} = -\frac{QS}{3}\begin{Bmatrix}1\\1\\1\end{Bmatrix}。$$

于是 $\bar{u}_i$ 可解出。

还应嵌入本质边界条件，但对于一个局部元素的解我们并不嵌入，这主要为总体元素求解作准备。

三、二次分布的三角形插值函数

在三角形每一条边上有三个节点，整个三角形有 6 个节点，如图 5.11 所示。其中 2、4、6 节点在各边的中点上，其基函数为：

$$\phi = a + bL_1 + cL_2 + dL_1^2 + eL_2^2 + fL_1L_2$$

根据每一节点上已知的 L_1, L_2 值，可列出 6 个方程，解出 a,b,c,d,e,f，得：

$\phi_1 = -L_1 + 2L_1^2$，　$\phi_3 = -L_2 + 2L_2^2$，　$\phi_5 = -L_3 + 2L_3^2$

$\phi_2 = 4L_1L_2$，　$\phi_4 = 4L_3L_2$，　$\phi_6 = 4L_3L_1$

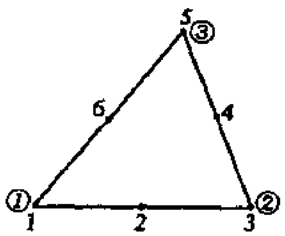


图 5.11

对 $\bar{L}(u) - p = 0$，照样可代入求出 A'_{NM}, F'_M，只是麻烦多了，方法类同于上面，这里就不介绍了。

四、三角形元总体有限元求解

问题：两块等边三角形平行平板对称放置，间距 $h \ll$ 边长，中间有流体存在，当二平板以等速 v 运动，不可压流体由其中挤出，二板中有压力场，其控制方程是：

$$\Delta p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = \frac{12\mu v}{h^3} = \mathrm{const} = G, p = p(x,y)$$

边界条件：$y = 0, y = 2 - \sqrt{3}x, y = \sqrt{3}x$ 三条边上，$p = 0$。

求解过程如下：

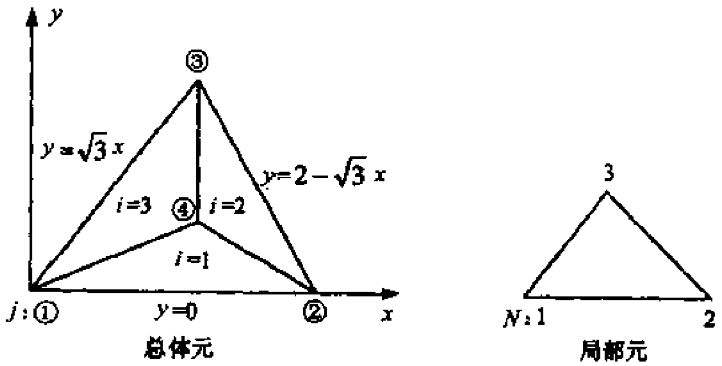


图 5.12

(1) 选元素

见图 5.12，总体元 4 个节点，3 个元素，$I=3$，$J=4$；局部元 3 个节点，$N=3$；在局部元与总体元节点耦合时，注意转向一致（逆时针），其中 x_4，y_4 为已知。

(2) 总体有限元方程

$$(A_{jk})(\bar{u}_j)=F_k \tag{5.51}$$

$$A_{jk}=\sum_{i=1}^{3}A_{NM}^{i}\Delta_{Nj}^{i}\Delta_{Mk}^{i} \tag{5.58}$$

$$F_k=\sum_{i=1}^{3}F_M^i\Delta_{Mk}^{i} \tag{5.59}$$

A_{jk} 为 4×4 矩阵，F_k 为 4×1 列矢量，由于 A_{NM}^i，F_M^i 由(5.74)与(5.74a)求出最后得：

$$A_{jk}=\begin{bmatrix} a_{11}^1+a_{22}^3 & a_{12}^1 & a_{21}^3 & a_{13}^1+a_{23}^3 \\ a_{21}^1 & a_{22}^1+a_{11}^2 & a_{12}^2 & a_{23}^1+a_{13}^2 \\ a_{12}^3 & a_{21}^2 & a_{22}^2+a_{11}^2 & a_{23}^2+a_{13}^2 \\ a_{31}^1+a_{32}^3 & a_{31}^2+a_{32}^1 & a_{32}^2+a_{31}^3 & a_{33}^1+a_{33}^2+a_{33}^3 \end{bmatrix}$$

四个方程解四个未知数，$(A_{jk})(\bar{u}_j)=F_k$，本质边界条件 $\bar{u}_1=\bar{u}_2=\bar{u}_3=0$ 嵌入：

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a_{31}^1+a_{32}^3 & a_{31}^2+a_{32}^1 & a_{32}^2+a_{31}^3 & a_{33}^1+a_{33}^2+a_{33}^3 \end{bmatrix}\begin{bmatrix}\bar{u}_1\\ \bar{u}_2\\ \bar{u}_3\\ \bar{u}_4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\S_1+S_2+S_3\end{bmatrix}\left(\frac{-G}{3}\right)$$

局部元素中各 x_i，y_i 均已知。

$$\begin{aligned}\bar{u}_4&=-\frac{G}{3}\,\frac{S_1+S_2+S_3}{a_{33}^1+a_{33}^2+a_{33}^3}\\ &=-\left(\frac{G}{12}\right)\frac{S_1+S_2+S_3}{\frac{1}{S_1}\left(b_3^{1^2}+a_3^{1^2}\right)+\frac{1}{S_2}\left(b_3^{2^2}+a_3^{2^2}\right)+\frac{1}{S_3}\left(b_3^{3^2}+a_3^{3^2}\right)}\end{aligned} \tag{5.75}$$

其中 S_i 为三个小三角形的面积，而 $b_3^1=y_1^1-y_2^1$ 等(5.74)有说明，将上式化简，最后得到：

$$\bar{u}_4(x_4,y_4)=-\frac{G}{3}\,\frac{(y_4-2+\sqrt{3}x_4)(y_4-\sqrt{3}x_4)y_4}{2(y_4+\sqrt{3}x_4)-3(x_4^2+y_4^2)}=p_4$$

而精确解为：$p=-\frac{G}{4}(y+2+\sqrt{3}x)(y-\sqrt{3}x)y$，令 $x_4=\frac{1}{\sqrt{3}}, y_4=\frac{1}{3}$，代入(5.75)所得的值与 $\bar{u}_4$ 相等。

三角形中心处的 p 与精确解一致，其它地方的 $\bar{u}_4$ 均有不同程度的误差，主要是总体元分成的局部元较少引起的。这里只是介绍求解的方法与过程，真正求解时应多分几个局部元。

整个过程以 Poisson 方程为例，对其它方程也可如法炮制，到此我们介绍了用迦辽金加权余量法总体元求解的全过程。

5.5 流体力学有限元法综述

主要介绍有限元法，尤其是迦辽金加权余量法在下述三种问题上的应用：

(1) 二维无粘不可压定常流；

(2) 二维粘性不可压流；

(3) 二维可压流。

5.5.1 二维不可压位势流

流体力学中介绍过二维不可压位势流的控制方程是 Laplace 方程：

$$\nabla^2\phi=0 \tag{5.76}$$

$$\nabla^2\Psi=0 \tag{5.77}$$

ϕ 为速势函数，Ψ 为流函数，其中 Ψ 给出第一类边界条件，即物面中一条流线，$\Psi_B=$ const；而 ϕ 给出第二类边界条件，即 $v_n=\left(\frac{\partial\phi}{\partial n}\right)_B=0$。$\phi$ 或 Ψ 求得后，速度即可解出，再算出压力 p，问题就解决了。

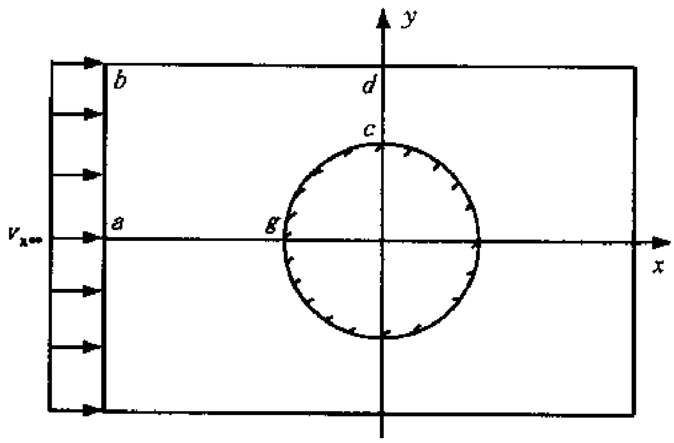


图 5.13

设无因次的来流 $u_\infty=1$ 流过 $R=1$ 的圆柱(见图 5.13)。这是一个上下左右对称的流场，只需分析 $abcdg$ 边界就可以了。边界上的 Ψ：

(1) agc 处 $\Psi = 0$，是零流线；

(2) ab 处：

$$\int_{\Psi_a}^{\Psi_b} \mathrm{d}\Psi = \int_{y_a}^{y_b} u_\infty \mathrm{d}y$$

$$\Psi = \Psi_a + u_\infty(y - y_a) = y$$

$$u_\infty(y_b - y_a) = \Psi_b = y_b$$

(3) bd 处：$\Psi = \Psi_b$；

(4) cd 处：$v = \dfrac{\partial \Psi}{\partial x} = 0$，$\left.\dfrac{\partial \Psi}{\partial x}\right|_{cd} = 0$

求解过程如下：

(1) 将网格划分成三角形元素，取四分之一流场，顺同一方向编号，预估物理量变化剧烈地方网格加密；

(2) 局部有限元求解 $\overline{\Psi} = L_N^i \overline{\Psi}_N^i$；

(3) 总体元求解：

$A_{jk} \cdot \overline{\Psi}_j = F_k = 0$，因为 $Q = 0$。其中 $A_{jk} = \sum\limits_{i=1}^{I} A_{NM}^i \Delta_{Nj}^i \Delta_{Mk}^i$。

对于我们的问题可取 $I = 111$，由于是 Laplace 方程，可以引用(5.74)式：

$$A_{NM}^i = \frac{1}{4S_i}\left\{\begin{bmatrix} b_1^2 & b_1 b_2 & b_1 b_3 \\ b_1 b_2 & b_2^2 & b_2 b_3 \\ b_1 b_3 & b_2 b_3 & b_3^2 \end{bmatrix} + \begin{bmatrix} a_1^2 & a_1 a_2 & a_1 a_3 \\ a_1 a_2 & a_2^2 & a_2 a_3 \\ a_1 a_3 & a_2 a_3 & a_3^2 \end{bmatrix}\right\}$$

将上述结果按布尔矩阵的划分原则进行计算，嵌入上述边界条件，(5.75)式只算了 $I = 3$，且边界条件也不同。现在 $I = 111$，由(5.51)，(5.58)，(5.59)式可借助计算机求解，我们假定已经完成。Δ_{Nj}^i，Δ_{Mk}^i 是布尔矩阵，其值由总体节点及局部节点的排列确定，从计算方法来说已解决了。

5.5.2 不可压粘性流动

$$\begin{cases} \dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} = 0 & (5.78\text{a}) \\ u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial y} = -\dfrac{1}{\rho}\dfrac{\partial p}{\partial x} + \upsilon\nabla^2 u & (5.78\text{b}) \\ u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} = -\dfrac{1}{\rho}\dfrac{\partial p}{\partial y} + \upsilon\nabla^2 v & (5.78\text{c}) \end{cases}$$

求解方法有下面三种。

一、流函数法

$$\Psi_y = u, \quad \Psi_x = -v$$

$$\frac{\partial}{\partial y}(5.78\text{b}) - \frac{\partial}{\partial x}(5.78\text{c})$$

消去压力项得：

$$\upsilon\nabla^4\Psi + \Psi_x\nabla^2(\Psi_y) - \Psi_y\nabla^2(\Psi_x) = 0 \tag{5.79}$$

这是非线性与非自伴的高阶偏微分方程，求解十分麻烦，即使用迦辽金法也要多次迭代。现设在 Re ≪ 1 蠕动情况下，对(5.79) 进行简化。方程中的 υ 是$\frac{\mu}{\rho}$，当 Re ≪ 1，即 $\upsilon \gg 1$，(5.79)式的后两项可略去，得：

$$\nabla^4 \Psi = 0 \tag{5.80}$$

迦辽金法分部积分：

$$\iint\left(\frac{\partial^2\Psi}{\partial x^2}\frac{\partial^2 L_N}{\partial x^2} + 2\frac{\partial^2\Psi}{\partial x\partial y}\frac{\partial^2 L_N}{\partial x\partial y} + \frac{\partial^2\Psi}{\partial y^2}\frac{\partial^2 L_N}{\partial y^2}\right)\mathrm{d}x\mathrm{d}y + \int_l \frac{\partial}{\partial n}\left(\frac{\partial^2\Psi}{\partial n^2} + \frac{\partial^2\Psi}{\partial s^2}\right)L_N\mathrm{d}l - \int_l \frac{\partial}{\partial n}\Psi_n \frac{\partial L_N}{\partial s}\mathrm{d}l = 0$$

首先 L_N 不能用线性插值 Lagrange 形式，必须保证节点处主要是交点处一阶导数连续，用 Hermite 插值函数为基函数即可以保证节点处流函数及其一阶导数连续。式中的 ∂n 与 ∂s 分别指边界曲线的法向与切向导数。有了基函数后其余作法与的前面一样，首先求局部元的 A^i_{NM} 与 F^i_M，再得 A_{jk}，F_k，然后加入本质边界条件求解。

二、流函数涡量法

上述方程化为：

$$\begin{cases}\omega = \dfrac{\partial v}{\partial x} - \dfrac{\partial u}{\partial y}\\ \Psi_y\omega_x - \Psi_x\omega_y = \upsilon(\omega_{xx} + \omega_{yy})\\ \Psi_{xx} + \Psi_{yy} = -\omega\end{cases} \tag{5.81}$$

令 $\overline{\Psi} = \Psi_N L_N$，$\overline{\omega} = \omega_N L_N$，用迦辽金法求解上述非线性方程组的主要困难是边界条件不好提，$\overline{\omega}$ 的边界条件不易提，因此不易求解。

三、解联立的 N.S. 方程：速度压力法求解，直接联立求解

(1) 迦辽金方法分别处理各基本方程，联立解代数方程组。

(2) 为使方程线化，惯性项前的变量用上一次计算的结果或近似解代入。令

$$\begin{cases}\overline{u} = \overline{u}_N\phi_{1N}\\ \overline{v} = \overline{v}_N\phi_{2N}\\ \overline{p} = \overline{p}_3\phi_{3G}\end{cases}$$

质量方程用压强的基函数来加权是为了便于方程联立求解。

$$\iint_\Omega \frac{1}{\rho}\phi_{3G}\left(\frac{\partial\overline{u}}{\partial x} + \frac{\partial\overline{v}}{\partial x}\right)\mathrm{d}\Omega = 0$$

$$\iint_\Omega \phi_{1N}\left(u_n\frac{\partial\overline{u}}{\partial x} + v_n\frac{\partial\overline{u}}{\partial y} + \frac{1}{\rho}\frac{\partial\overline{p}}{\partial x} - \upsilon\nabla^2\overline{u}\right)\mathrm{d}\Omega = 0$$

$$\iint_\Omega \phi_{2N}\left(u_n\frac{\partial\overline{v}}{\partial x} + v_n\frac{\partial\overline{v}}{\partial y} + \frac{1}{\rho}\frac{\partial\overline{p}}{\partial y} - \upsilon\nabla^2\overline{v}\right)\mathrm{d}\Omega = 0$$

其中 u_n，v_n 是上一次计算的结果或近似解的结果的当地值，可视为常数，由于最高阶次一般是一阶的，即 $\overline{u}$，$\overline{v} \in w_2^1$，粘性项可以分部积分，因此可用 Lagrange 线性插值函数作为基函数，如用三角形元可令 $\phi_{1N} = \phi_{2N} = L_N$，质量方程：

$$\iint_\Omega \frac{1}{\rho}\phi_{3G}\frac{\partial L_N}{\partial x}\mathrm{d}\Omega\{\overline{u}_N\} + \iint_\Omega \frac{1}{\rho}\phi_{3G}\frac{\partial L_N}{\partial y}\mathrm{d}\Omega\{\overline{v}_N\} = 0$$

x,y 向动量方程也可类似写出 。

1974 年 Hood. p 与 Tayer. G 在解 N. S. 方程中就提出最好使速度的基函数取二次分布 Lagrange 插值函数，而压强的基函数则用线性分布的 Lagrange 插值函数为基函数，得到的速度与压强的精度都是一阶。如果速度与压强都用一种基函数表示，那么速度精度高而压强较差，解不太合理。

另外，边界条件可由两部分组成，一部分给出当地的速度值，一部分给出边界上的压力值。速度压力法的优点：

(1) 易推广到三维，流函数法和涡量法不行；

(2) 只要保证交点节点处变量存在，可用 Lagrange 插值函数作为基函数；

(3) 速度压力边界条件直接嵌入矩阵求解；

(4) 计算时间较少。

5.5.3 可压粘性流动

以二维为例，简化的 P.N.S. 方程为(1.40)式：

$$\begin{cases} u\dfrac{\partial\rho}{\partial x}+v\dfrac{\partial\rho}{\partial y}+\rho\left(\dfrac{\partial u}{\partial x}+\dfrac{\partial v}{\partial y}\right)=0 \\ uu_x+vu_y+\dfrac{p_x}{\rho}=\dfrac{v}{Re_*}u_{yy} \\ uv_x+vv_y+\dfrac{p_y}{\rho}=\dfrac{4v}{3Re_*}v_{yy} \\ \rho\left(u\dfrac{\partial T}{\partial x}+v\dfrac{\partial T}{\partial y}\right)-\left(u\dfrac{\partial\rho}{\partial x}+v\dfrac{\partial\rho}{\partial y}\right)N=\dfrac{\mu}{Re_*Rr}\dfrac{\partial^2T}{\partial y^2}+\dfrac{N\mu}{Re_*}\left[\dfrac{4}{3}\left(\dfrac{\partial v}{\partial y}\right)^2+\left(\dfrac{\partial u}{\partial y}\right)^2\right] \\ p=\rho RT \end{cases}$$

近似解的表达式为：

$$\bar{\rho}(x,y)=\phi_N^\rho(x,y)\bar{\rho}_N$$
$$\bar{u}(x,y)=\phi_N^u(x,y)\bar{u}_N$$
$$\bar{v}(x,y)=\phi_N^v(x,y)\bar{v}_N$$
$$\bar{T}(x,y)=\phi_N^T(x,y)\bar{T}_N$$

当用迦辽金法处理积分关系式时，要注意选择适当的基函数。经验表明，用变量 $\bar{\rho},\bar{u},\bar{v},\bar{T}$ 为基函数是合适的，即：

质量方程：$I_1=<\varepsilon_i^1,\bar{\rho}>=<\varepsilon_i^1,\phi_N^\rho\bar{\rho}_N>=0$

动量方程：$I_2=<\varepsilon_i^2,\bar{u}>=<\varepsilon_i^2,\phi_N^u\bar{u}_N>=0$

$I_3=<\varepsilon_i^3,\bar{v}>=<\varepsilon_i^3,\phi_N^T\bar{v}_N>=0$

能量方程：$I_4=<\varepsilon_i^4,\bar{T}>=<\varepsilon_i^4,\phi_N^T\bar{T}_N>=0$ 按上述方程联立求解，解出 $\bar{\rho},\bar{u},\bar{v},\bar{T}$.

第六章　并行计算的基本概念

6.1 引言

计算流体力学(CFD)需要高速度、大容量的计算机。即使在目前工艺水平所能提供的浮点运算速度最快、容量最大的超级计算机上进行计算，一个三维非定常问题的数值模拟也要花费几十个CPU小时。用计算机对整架飞机进行设计，必须考虑湍流和真实气体效应，求解复杂的三维内流场和外流场，需要Teraflops级甚至Petaflops级的超级并行计算机。正是各个领域类似的大型科学计算问题，如核模拟、数值天气预报、石油地震数据处理、遥测遥感、粒子物理以及C^3I数据融合等，提出了对计算机性能越来越高的要求，促使计算机不断更新换代满足这种需要，才推动了计算机学科的不断向前发展。反过来，计算机学科的发展又促使计算流体动力学等需要高速度、大容量的计算机作为研究工具的学科和专业朝更深、更高、更广的方向拓展，进而又提出对计算机性能更高的要求，成为设计新一代计算机的源动力。

自从1946年第一台电子数字计算机问世以来，计算机的发展先后经历了四代，目前正向第五代过渡。人们通常习惯以电子管(V)、晶体管(T)、中小规模集成电路/大规模集成电路(LSI)、超大规模集成电路(VLSI)等器件的变革作为计算机换代的标志。当然，这种器件的更新，其速度、功能、可靠性的不断提高和价格的不断降低，确实在不断地推动着计算机的发展，是计算机性能得以提高的物质基础。但是，如果把某一时期器件性能的提高和同一时期计算机系统性能的提高作一番比较，不难发现：计算机系统性能提高的指数要比器件性能提高的指数大得多，其中计算机体系结构的改进是计算机系统性能得以提高的重要因素。这种情况在计算机近年来的发展中变得更加明显，这是因为器件性能的提高受门切换和光传输速度的限制，已经越来越接近它所能达到的极限，每一点点提高都要求在器件的制作水平与工艺上付出成倍的努力。解决上述问题的出路只有一条，不断更新计算机的体系结构，其突出特点就是将并行性越来越多地引入到计算机系统的设计中。

6.2 基本概念

为以后叙述方便起见，我们先介绍与并行计算有关的一些基本概念和术语。

(1) 时间与工作量

对于在独占方式下的单CPU系统，CPU时间经常用来度量所完成的工作量，此时CPU时间和墙上时间(Walltime)比较接近。多个CPU同时运行的能力使得工作量和完成该工作所用的时间有必要加以区分。在多CPU环境下运行多任务或多进程程序，可以在更短的墙上时间内完成单个CPU所完成的工作，但由于存在各类开销，程序运行的CPU

时间的总和一般会有所增加。

(2) 粒度(Granularity)

通常指模拟问题的计算规模，但近年来也有人将它定义为模拟问题的计算量与通信量的比值。

(3) 加速比(Speedup)

相对最优的串行程序在单 CPU 上执行的墙上时间与模拟同一问题的并行程序在多 CPU 上执行的墙上时间之比。

(4) 效率(Efficiency)

加速比与并行计算所用 CPU 数之比。

(5) 开销(Overhead)

为了保证运行结果的正确性，在并行程序设计中可能要引进一些冗余计算，进行必要的同步和通信，导致并行计算的工作量与串行计算相比会有所增加，增加部分所花费的时间就是所谓开销。各处理机的开销记为 Q_i，总开销记为 Q_t，显然有 $Q_t = \sum_i Q_i$。记当某处理机并行计算的墙上时间与 T_p 完全一致时的开销为 Q，通常情况下可用 Q_t/p 近似表示 Q。

(6) 性能价格比

算法在机器上执行时所取得的性能与机器价格的比值，可用于指导用户在购买机器时，在性能与价格之间折衷。

(7) 代价(Cost)

并行程序在多 CPU 上执行的墙上时间与处理机台数的乘积。

(8) 并行开销(Parallel Overhead)

代价与同一模拟问题的串行程序在单 CPU 上执行的墙上时间之差。

(9) 并发度 C

在算法中能同时执行的最大任务数。

(10) 冗余度 $R(p)$

p 台处理机完成的总的单元操作数(非发呆时间的总和)与单台处理机完成同等规模问题的单元操作数之比，表示软件并行性与硬件并行性间的匹配程度。

(11) 利用率

效率与冗余度的乘积，经化简可知，它表示单位时间内在各处理机上并行执行所完成的平均操作数与在单机上串行执行所完成的操作数之比，即并行程序在执行过程中资源(处理机、存储器等)保持在忙碌状态下的百分比。

(12) 并行性质量的计算公式为：

$$Q(p) = \frac{S(p)E(p)}{R(p)} = \frac{T_1^2 N(1)}{pT_p^2(p)N(p)}$$

其中 $N(p)$ 表示在 p 台处理机上完成的总的单元操作数，$N(1)$ 表示在单台处理机上完成同等规模问题的单元操作数。显然，加速比越高，并行效率越高，并行质量越好；冗余度越大，并行性质量越差。

(13) 同步(Synchronization)

同步是协调各并行任务或进程的一种方法，它保证相互作用的任务或进程所要求的

条件得到满足。同步点是相互合作的任务或进程协调各自运行状态的地方。

(14) 死锁(Deadlock)

由于错误地使用同步机制,使得一个任务或进程等待一个永远不可能发生的事件,就称为死锁。经常碰到的死锁形式是两个任务或进程互相等待对方完成某个动作。

(15) 确定性(Determinism)

在一个CPU上运行的程序,其语句的执行次序是良好定义的。因为每次运行,相同的指令以相同的次序执行,所以重复的运行产生相同的结果(包括中间结果)。数据和控制相关简单地通过语句在程序中的位置得到满足,执行的次序由语句在程序中的次序确定。多任务或多进程对执行次序引入了一种新的因素,并行意味着不确定,不确定意味着不可重复,虽然在每一个任务或进程中,语句的执行次序仍然是良好定义的,但其相对次序可以在每一次运行都不同,运行错误可能是不可重现的,因而,并行程序更难于测试和调试,传统的调试方法,如在程序中设置打印语句查错的强力纠错等方法已不能满足并行程序调试的要求,而目前可用来帮助分析和防止这类错误的工具不多。

(16) 负载平衡(Load Balance)

负载平衡是一种使执行程序的多个CPU各自完成大致相当的工作量的技术,分为静态负载平衡和动态负载平衡两类。静态负载平衡是各任务或进程所要完成的工作在运行前就确定了的负载划分方法,动态负载平衡是各任务所完成的工作要在运行时才可确定的一种负载划分方法,用户应利用并行系统软件提供的特殊支持,使已完成任务的CPU不断地获取下一份未完成的工作并完成它,直到全部工作完成。

6.3 并行性与并行性等级

所谓并行性(Parallelism)是指在同一时刻或是同一时间间隔内完成两种或两种以上性质相同或不相同的工作。只要时间上互相重叠就存在并行性。严格来讲,把两个或多个事件在同一时刻发生的并行性叫做同时性(Simultaneity);而把两个或多个事件在同一时间间隔内发生的并行性叫做并发性(Concurrency)。

并行性可以有不同的等级。从执行程序的角度看,并行性等级从低到高可以分为:

① 指令内部并行——指令内部的微操作之间的并行;

② 指令间并行——并行执行多条指令;

③ 作业级或过程级并行——并行执行多个进程或任务;

④ 作业或程序级并行——在多个作业或程序间的并行。

在计算机系统中,可以采用时间重迭,资源重复和资源共享等多种技术途径来提高并行性。

6.4 计算机分类

计算机可以按不同的原则进行分类,其中按指令流和数据流的性质来划分的弗林(Flynn)分类法比较直观,易于接受。为叙述方便先给出如下术语:

① 指令流(Instruction stream)——机器执行的指令序列。

② 数据流(Data stream)——由指令流调用的数据序列包括输入数据和中间结果。

③ 多倍性(Multiplicity)——在系统最受限制的元件上,同时处于同一执行阶段的指令或数据的最大可能的个数。

按指令流和数据流的多倍性,可将计算机系统分为下列四类:

- 单指令流单数据流(SISD)计算机;
- 单指令流多数据流(SIMD)计算机;
- 多指令流单数据流(MISD)计算机;
- 多指令流多数据流(MIMD)计算机。

SISD 计算机就是传统的串行计算机或 Von Neumann 计算机。在这类计算机中,也可以融入并行处理技术,如多功能部件、流水线、超高速缓存、多级存储、超长指令字、多线程、超标量以及向量指令与链结等。但是,只要它的指令部件一次只对一个指令进行译码并且只对一个操作部件分配数据,则它仍属于 SISD 系统。各类工作站和 PC 机均属此类,如 IBM390,VAX,Alpha,Sparc 等。

SIMD 计算机的一个典型代表是阵列式处理机。这一类计算机的主要特点是各个处理单元的指令来自中央处理部件,每一个处理单元在同一时刻执行同一指令,但使用不同的操作数据。如阵列机 STAR100 和 DAP、流水线向量计算机 Cray-1 和 YH-1、分布存储并行计算机 CM-2 均可划入 SIMD 类。

MISD 计算机由多个处理机组成,每一个处理机都有自己的控制部件,它们共享一个公共存储器,它有多条指令流和一条数据流。在执行的每一步,每台处理机从它的控制部件接收指令,并从共享存储器接收同一数据进行操作。目前 MISD 系统代表何种计算机尚有争议,实际上也很少有人对 MISD 系统的结构和适用性进行研究。

MIMD 计算机是指各处理机在同一时刻执行不同的指令,指令来自处理机本身,各处理机不一定要同步运算。复杂的程序被分成若干可独立运行的程序,在不同的处理机上分别执行并相互交换信息,从而实现了各个级别的全面并行性。共享内存的多处理机系统 Cray X-MP、Cray Y-MP、Fujitsu VPP500、YH-2 以及大规模并行计算机 IBM SP2、Intel Paragon、nCUBE、CM-5、YH-3、各种 SMP 和 Cluster 系统等均可划入 MIMD 类。

6.5 并行加速比

一、Amdahl 定律与 Gustafson 修正及其等价性

如 6.2 节所述,一个并行计算程序的加速比 S_p 定义为相对最优的串行程序在单 CPU 上执行的墙上时间与模拟同一问题的并行程序在多 CPU 上执行的墙上时间之比。设 T_1 是相对最优的串行程序在单 CPU 上执行的墙上时间,T_p 是模拟同一问题的并行程序在 p 台处理机上执行的墙上时间。T_s 为程序中必须顺序执行的部分,T_{para} 为程序中可以在 p 台处理机上并行执行的部分,α 为程序的串行部分占总执行时间的百分比,即 $\alpha = T_s/T_1$,相应地,$1-\alpha$ 为多任务或多进程部分所占原始执行时间的百分比,称为并行化率。O 为开销(Overhead)。从而并行程序的加速比为 $S_p = T_1/T_p = T_1/(T_s + T_{para}/p + O)$。忽略各类开

销，则有

$$\begin{aligned}S_p(p,\alpha)&=T_1/(T_s+T_{para}/p)\\&=T_1/[\alpha T_1+(1-\alpha)T_1/p]\\&=1/(\alpha+(1-\alpha)/p)\\&=p/(1-\alpha+\alpha p)\end{aligned}\tag{6.1}$$

这就是 Amdahl 法则。它表明：

（1）除非程序的绝大部分都并行化，否则不可能达到高的加速比。比如，如果在4CPU 上有 50%并行化代码，可能达到的最大加速比只有 1.6。

（2）对于固定的 α 值，p 变化时，加速比的增长不如 p 的增长快。p 趋于无穷时，加速比收敛于一个有穷值。对于具有许多处理机的系统，程序执行的时间主要由非任务部分决定，此时并行化率越高越好。

按照 Amdahl 定律，即使串行比例 α 只占 0.1，用 1000 个处理机的最大加速比不会超过 10。然而 1988 年 2 月美国 Sandia 国家实验室的 Gustafson 等三人在由 1024 个处理机以超立方体网络互连的 ncube 系统上进行了 3 个应用问题的试算，并行加速比分别达到 1016、1020 与 1021。依据这些试算结果，他们指出，经典的 Amdahl 定律蕴涵着问题规模不变的假设，即算法中的并行成分不随处理机台数变化的假设，是不合实际的。事实上，求解问题的规模随着处理机增多而扩大，并行度也随之增加。他们指出，应假定串行部分所耗的时间为常数，设在并行程序中只能串行执行的运算量的百分比为 α，那么在并行程序中，串行部分执行的时间保持不变，仍为 αT_p，但并行部分串行做的时间就需要 $p(1-\alpha)T_p$，故并行加速比为

$$S_p=\alpha+(1-\alpha)p\tag{6.2}$$

此式称为加速比定律的 Gustafson 修正。Amdahl 定律是用单处理机能串行求解的小规模问题去测定并行机的加速比，而 Gustafson 修正则是从并行机所能求解的大规模问题的性能去度量它与串行机的加速比，认为 α 不是一成不变的，而是随着处理机台数的增加和所求解的问题规模的扩大而不断减小的。

事实上，用 Gustafson 修正中的 α，即并行程序中只能串行执行的百分比来表示串行程序中只能串行执行的百分比 β，则 $\beta=\alpha/(\alpha+(1-\alpha)p)$，再按 Amdahl 定律求加速比

$$\begin{aligned}S_p&=\frac{p}{1-\beta+\beta\times p}\\&=\frac{p}{1-\dfrac{\alpha}{\alpha+(1-\alpha)p}+\dfrac{\alpha}{\alpha+(1-\alpha)p}\times p}\\&=\alpha+(1-\alpha)p\end{aligned}\tag{6.3}$$

这就是 Gustafson 修正，而它却可以通过 Amdahl 定律求出，可见二者本质上是一致的。

二、超线性加速比

由 Amdahl 定律有

$$S_p-p=\frac{p}{1-\alpha+\alpha\times p}-p=\frac{\alpha(1-p)p}{1-\alpha+\alpha\times p}\leqslant 0\tag{6.4}$$

由 Gustafson 修正有

$$S_p - p = \alpha + (1-\alpha)p - p = \alpha(1-p) \leqslant 0 \tag{6.5}$$

故均不大可能出现超线性加速比即 $S_p > p$ 的情况。但是在某些分而治之的算法中，如果并行运算步小于相应的串行运行的总运算步，则在负载平衡性很好、各类开销很小的情况下可能出现超线性加速比。例如，并行搜索允许每个处理机在不同的方向上搜索，一旦某个进程找到了所要的结果，它就要向其它处理机发信号以终止搜索，这样就有可能比相应的串行算法提前取消那些无谓的搜索分支，出现超线性加速比。但在一般情况下，当并行出现超线性加速比时，要很慎重地分析原因，并行计算机系统中各个处理机的存储方式的不同，高速缓存(Cache)的影响，以及并行操作系统开销的均摊都有可能出现超线性加速比的假象。如果并行计算程序不是从串行计算程序改造过来的，而是直接在并行机上调试开发出来的，那么并行加速比可定义为求解某问题的并行计算程序在单处理机上运行的墙上时间 T_1 与该程序在 p 台处理机上运行的墙上时间 T_p 之比。这样得到的并行加速比通常要略大于原始意义上的并行加速比，这是因为并行操作系统的开销一般大于串行操作系统的开销。

三、几种求 T_1 的变通方法

在分布式并行计算中，容易出现的一个问题是超内存而根本无法直接测量在单处理机上串行计算究竟要花多少时间，这样在实际计算加速比时有如下几种求 T_1 的变通方法：

(1) 将各处理机运行的 CPU 时间全部加起来作为 T_1，这样得到的加速比不大合理，没有加速也会计算出加速，反过来，出现超线性也计算不出超线性，即使对并行搜索算法也是如此，故一般不宜采用。

(2) 对串行程序的工作量进行量级分析(当规模较小时不能扔掉所谓的低阶项)，然后理论预测当规模达到并行计算程序目前所模拟的规模时的 T_1，不便之处是某些实际计算问题根本无从进行理论上的量级分析。

(3) 根据一系列单处理机小规模串行计算的运行时间用最小二乘法拟合一条计算量与规模关系的经验曲线，求出 T_1。

(4) 程序在 k 台处理机上并行执行不超内存，但在 $k-1$ 台处理机上并行执行超内存，可用 $T_k \cdot k$ 作为 T_1，这样得到的 T_1 会大于原始意义上的 T_1。

6.6 并行计算模型研究

一、研究并行计算模型的动机

20 世纪 80 年代以来，出现了多种并行计算机，其体系结构不尽相同。为了取得较好的性能，算法设计需要涉及机器的体系结构与具体参数，这就阻碍了算法软件的商业化，同时不利于并行算法的研究及其软件移植。

为此，把机器的少量关键参数抽象出来，形成一类抽象的计算机结构，就称之为并行计算模型，成为算法设计的一个基本界面，并行计算模型的引入具有十分重要的作用，主要表现在：

(1) 具体的计算机多种多样，面对具体机器，设计与机器参数相关的算法是相当繁杂

的工作，也不利于算法的研究及其软件移植，影响算法软件的标准化和商业化。引入并行计算模型后，只要设计的并行算法是针对该计算模型的，且模型考虑到了这样的结构，则设计的算法可移植到该计算机上，而细节工作可交给程序员去做。

（2）可利用并行计算模型去设计、分析并行算法，进而去预测算法和程序的执行时间，指导机器设计。

（3）由于避开了各种各样的具体并行机结构，算法研制者可借助并行计算模型去设计算法，集中开发与应用问题相关的并行性，并分析算法性能，从而使并行算法研究成为一个专门的研究领域。

（4）指导用户购买机器。针对具体应用领域，通过分析得到该领域的问题对哪些参数（模型参数）依赖性强，而去购买相关机器参数较好的机器。

二、PRAM 并行计算模型

该模型假设系统具有一个容量无限大的共享存储器，并且通信带宽没有任何限制，具有有限台或无限台完全同构的处理机。PRAM 模型按控制方式的不同分为两类，即 SIMD PRAM 模型与 MIMD PRAM 模型。

1. SIMD PRAM 模型

在本模型中，任一处理机可以通过 SM 同任一其它处理机交换数据，这种模型按对同一共享存储单元的存取访问的约束条件进一步分为：

· SIMD-EREW PRAM 模型，不允许同时读或写同一共享单元；

· SIMD-CREW PRAM 模型，允许同时读同一共享存储单元，但不允许同时写；

· SIMD-CRCW PRAM 模型，允许对同一共享单元同时读及同时写，同时写可能会引起数据的不一致性，因而按写的方式不同又分成为三种类型：允许具有同样值的机器同时写，从要求同时写的机器中任取出一台写共享单元和从要求同时写的机器中找出优先级最高的处理机去写。

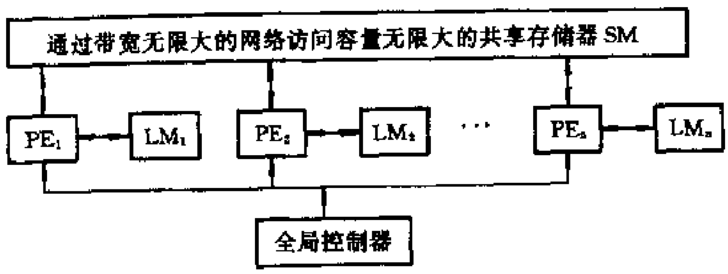


图 6.1 SIMD PRAM 模型示意图

2. MIMD PRAM 模型

按互连网的不同可以进一步分为：

· 总线型即 MIMD-Bus-PRAM 模型；

· 交叉开关型即 MIMD-CB PRAM 模型；

· 多级互连网型即 MIMD-MIN PRAM 模型。

MIMD PRAM 模型的特点：

· 可多任务分布计算；
· 因为控制器的局部性，所以要用共享变量显式同步，开销大；
· 负载可能不平衡。

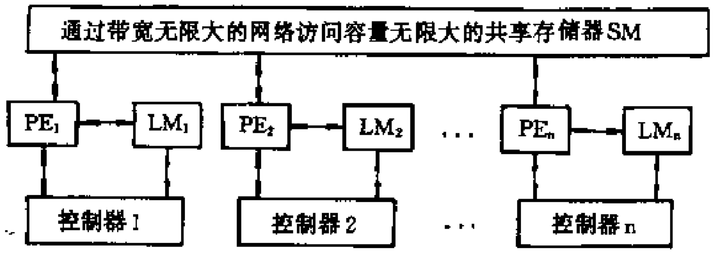


图 6.2 MIMD PRAM 模型示意图

3. PRAM 模型的局限性

· 假设都是不切实际的，如假定每台处理机在单位时间内都能访问 SM 中任意单元，假设通信带宽无限大，等待时间为 0(无阻塞)，无额外开销，各处理机完全同步地工作等等；

· hot-spot 问题，用锁机制则开销大，如人工安排程序中的语句次序来减小同步开销，则工作相当繁杂。

4. PRAM 模型的优点

在 PRAM 模型上设计算法简单，且易与传统的串行计算模型 RAM 模型进行联系。

5. PRAM 模型上的并行计算时间

设第 i 台处理机上的工作负载所用时间为 w_i，同步时间为 l，则总的计算时间为：

$$T = \max_i w_i + l \tag{6.6}$$

三、BDM 模型

BDM 模型为块分布存储模型。它是共享存储编程模式与基于消息传递的分布式存储系统模式之间的桥梁模式。它用四个参数描述：处理机数目 p、主存中的连续字 m、处理机从发出访问请求到得到远程数据的最大延迟时间 τ(包括准备请求的时间、请求包在网络中路由的时间、目的处理机接受请求的时间以及将包含在主存中的连续 m 个字返回给源处理机的时间）以及处理机发送数据到网络或网络接收数据的时间 σ。那么，对远程访问共需 $\tau + m\sigma$ 时间，k 个处理机访问同一主存段，需 $k(\tau + m\sigma)$ 时间完成。BDM 模型认可流水线预取技术，即某个处理机的 k 次预读操作可在 $\tau + km\sigma$ 时间内完成，通过数据访问进行消息传递。

BDM 模型的优点在于：

(1) 用 m 反映了空间局部性特点，提供了一种评价共享主存算法性能的方法，度量了因远程访问引起的处理机间的通信；

(2) 描述了流水线预取技术；

(3) 可编程性好；

(4) 考虑了共享主存中的存储竞争问题；

(5) 可以用来分析网络路由情况。

BDM 模型的缺点在于：

(1) 认为初始数据放置于局部存储器中，对于共享主存程序的编程者来说，需要额外增加数据移动操作；

(2) 对网络中影响延迟的因素，如处理机的本地性及网络中的拥挤等未加考虑；

(3) 未考虑系统开销。

四、QSM 模型

QSM 模型为排队共享存储模型。它只有两个参数，即处理机数和间隔参数。它可以反映本地指令速率与通信速率之间的差异；该模型使用了 QRQW PRAM 中的排队竞争原则，即每个存储单元都有一个队列与它相联系，每一个要求访问它的请求均使用 FIFO 进行排队。处理机每隔时间就发出一个访存请求，而该请求的服务时间为单位时间(不包括等待时间)。

这个模型描述了存储竞争，但是没有考虑空间局部性；它也没有考虑处理机上的存储层次。不论本地工作集的大小如何，均假定一个本地计算需要花费单位时间(即没有考虑处理机资源的竞争)；这个模型使用了一个带宽参数(即局部带宽)来反映每个处理机面临的带宽限制。但是贝尔实验室的 Micah Adler 等人发现，当对一些基本问题(如广播、求和、分类等)分别采用基于聚合式带宽和基于局部带宽的模型进行分析时，在条件相同的情况下，前者要比后者好一些。

五、BSP 模型

1. 引入 BSP 模型的动机

在基于消息传递模型所设计的软件中，为了没有等待开销而恰当地创建 send 和 recv 通信对将相对困难，而且由于网络带宽的限制，通信量较大时阻塞较严重，因而难以利用这种模型准确地预测通信延迟和计算性能。

共享主存模型易于编程，但要控制对同一共享单元的访问，要求程序编写者有高超的技巧或要用到昂贵的锁机制。实现共享主存模型的抽象需要使用大量计算机资源来处理通信和保持一致性。

在上面几种模型有其局限性的考虑之下，BSP(Bulk Synchronous Parallelism)模型问世。

2. 什么是 BSP

BSP 既是一种并行计算模型，又是一种程序设计方法，面向所有应用领域和广泛体系结构的通用并行性，BSP 具有以下基本特征：

· 程序编写容易。BSP 程序很像串行程序，只有极少的额外信息用来支持描述并行。

· BSP 与具体机器结构无关，可移植性好。

· 在给定机器上程序的性能是可预测的，BSP 程序的执行时间可从代码本身与少量几个机器参数算出。

· BSP 放弃了把局部性作为性能优化的标准，采用进程到处理机的随机映射。这简化了具体实现时的许多问题，但对某些问题，BSP 不是最佳选择。

3. 用 BSP 模型设计的程序的基本结构

BSP 程序由一个或多个 Superstep 组成，每个 Superstep 由本地计算、全局通信以及一个同步组成。

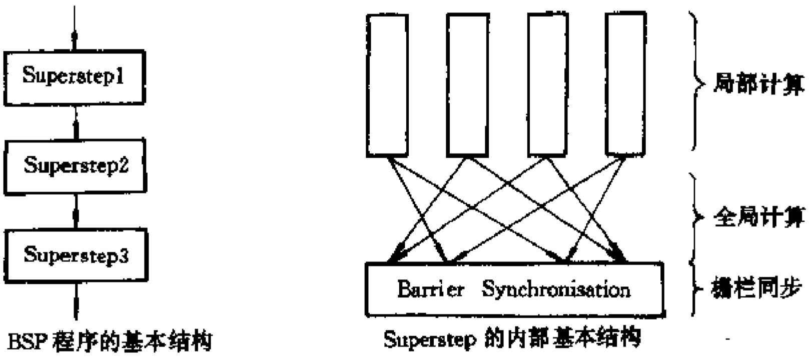


图 6.3

4. 通信作用的发挥

BSP 把一个 Superstep 看为一个单元，其间通信会有一些发送操作和接收操作，设发送量和接收量中最大者为 h，则称该 Superstep 中的通信为 h-relation。这时一个通信步相当于一个 permutation 通信。

通信网传输数据的能力用参数 g 来描述，它表示网中相邻步(时间上)的两个消息发送时的平均最小间隔。这样一个 h-relation 通信时间为 hg，g 常用 Benchmark 测得。

5. BSP 程序的运行时间

每个 Superstep 运行时间为：

$$t_{\text{Superstep}} = \max_i w_i + \max_i(h_i)g + l \tag{6.7}$$

其中 w_i 是该 Superstp 中 process i 的负载，h_i 是联系于进程 i 上的通信 hi-relation，而是每个栅栏同步的时间。由此可知，要在 BSP 模型上取得较好的性能，应该：

- 平衡负载；
- 平衡通信；
- 极小化 Superstep 的总数；
- 通信尽量后推到 Superstep 后端以减小通信起步次数和时间总量。

6. BSP 使用情况

BSP 并行平台 BSPLib 可从网站 http://www.bsp-worldwide.org/implements/ox-tools.htm 得到。BSP 模型已用于流体力学的计算和电磁学的计算等领域。

7. BSP 与 PRAM 模型的关系

$g=0$ 是 PRAM，$l=0$，$g=0$ 认为是理想的 PRAM 模型。

六、LogP 模型

1. LogP 模型的基本概念和参数刻画

LogP 是一种面向分布存储，点对点通信的多计算机模型。该模型充分说明了互连网络的基本性能特点，而且没有涉及网络的结构。模型主要由以下几个参数来刻画：

- L(Latency)：模块间进行通信(一个字或 n 个字)所要的等特时间或延迟时间的上

限。

• o(overhead)：处理机用于发送或接收每个消息的时间开销，在这段时间内，处理机不能执行别的操作。

• g(gap)：一台处理机连续地进行消息发送或连续地进行消息接收时的最少时间间隔，g 的例数与可用的每台处理机的通信带宽相对应。

• P(Processor)：处理机或存储模块的数量。

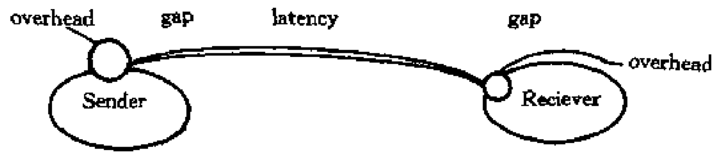


图 6.4

2. 消息传递模式

3. LogP 模型的局限性和特性

• LogP 模型是异步的，即各处理机是异步工作的，任何消息经历的等待时间不可预测。多个消息的发送次序与接收次序会有所不同(多信道情形)。

• 在 wormhole 技术下，L 可忽略不计(在长消息情况下)。

4. LogP 模型下程序的执行时间

$$T_{para} = \max_i w_i + (N-1)\max(o,g) + L + 2o \tag{6.8}$$

其中 w_i 是第 i 台机上的负载的运行时间，N 是消息总量(假设仅通信了一次，对通信多次时，起步时间再乘以次数 m，$N-1$ 换成 $N-m$ 即可)。

5. Log$P$ 模型与 BSP 模型的关系

$$\text{Log}P + \text{barriers} - \text{overhead} = \text{BSP}$$

七、C^3 模型

C^3 名字来源于三个英文单词“计算”(Computation)、“通信”(Communication)与“拥挤”(Congestion)的头三个字母。它有四个主要参数：l 表示消息包的长度，与带宽有关；s 为启动时间；h 表示延迟，常用两个处理机间的平均距离表示；p 为处理机数目。同 BSP 模型类似，它也认为并行计算由一个超步的序列组成。在对通信部分进行度量时，使用了两个量来描述网络的拥挤：连接上的拥挤量和处理机上的拥挤量；并且考虑了不同路由模式和不同发送、接收原语对通信的影响。但是，它用来度量连接上的拥挤量的参数有一个前提或假设，即同一个通信对中的两个处理机要处于对分网络后的不同子网中；模型假设了网络带宽等于处理机带宽，这影响了正确描述系统的可扩展性。该模型适用于粗粒度系统。

八、可能的发展方向

当前并行计算模型研究的可能发展方向有：

(1) 使用聚合式带宽或层次式带宽作为参数；

(2) 对多线程技术中上下文的开销进行度量；

(3) 共享存储机器中的通信竞争、非阻塞请求的模型化；

(4) 描述处理机与存储器之间的瓶颈，如处理机速度与访存速度不匹配等；

(5) 如有可能，考虑计算与通信重叠的问题。

6.7 并行系统可扩性分析

一、可扩性分析的定义

可扩性是描述问题规律增大时，系统用增加的处理机进行计算时的能力，是描述系统规模增大时的性能增长问题。它是机器和算法的组合体即并行系统的可扩性，单纯讲算法可扩性或机器可扩性都不恰当。

一般性理解，可扩性即 Scalability，含有比例化的意义，描述固定某些参数之后，另一些参数间是否会有按比例变化。

二、性能可扩性的基本特征

(1) 应提供系统规模如何影响性能的信息，即可扩性应为系统规模的函数。

(2) 应是描述算法与机器的组合体的函数，因为性能受到算法与机器两方面因素的影响。

(3) 应可评估，可对不同组合体进行比较，即应定量或定性到一定程度。

三、可扩性分析的作用

(1) 用于分析算法及其实现时的瓶颈，以及选择最优的算法以充分利用增加的处理机。

(2) 用于研究互连网的内在开销，以及其它硬件技术对性能的影响，从面更有利于理解科学计算的特性，以促进机器制造。

(3) 用于并行计算性能评价和预测。

(4) 用于选择最优的处理机数以取得某个最佳性能参数。

四、传统的可扩性模型介绍与理解

1. 问题规模受限模型——Amdahl 定律(加速比定律 1)

设问题规模不变，则总的负载不变，设

$$w = w_s + w_{para} \tag{6.9}$$

其中 w_s 是程序只能串行执行的部分，则

$$S_p = (w_s + w_{para})/(w_s + w_{para}/p) \tag{6.10}$$

设 $w_s = \alpha w$(恒定)，则 $w_{para} = (1-\alpha)w$，从而

$$S_p = \frac{1}{\alpha + (1-\alpha)/p} \to \frac{1}{\alpha} \quad (p \to \infty) \tag{6.11}$$

由此可知处理机无论如何增加，加速比将受到程序中串行部分所占比例的限制。显然，可以认为串行部分所占比例越小，可扩性越好；反之，则可扩性越差。这个可扩性模型主要涉及算法，而对机器的涉及相对较少。

2. 时间受限模型——Gustafson 修正(加速比定律 2)

设系统规模从单台增加到 p 时，相应地把负载从 $w^{(1)}$ 增加到 $w^{(p)}$，使得前后系统执行

时间一样。

假设 $w^{(p)} = w_s^{(p)} + w_{para}^{(p)}$，且 $w_s^{(p)} = w_s^{(1)}$，则由 $T_p = T_1$ 得：

$$w_s^{(1)} + w_{para}^{(1)} = w_s^{(p)} + w_{para}^{(p)}/p \tag{6.12}$$

其中 $w_s^{(p)}$ 是 $w^{(p)}$ 的串行部分，$w_s^{(1)}$ 是 $w^{(1)}$ 的串行部分，由此有：

$$w_{para}^{(p)} = pw_{para}^{(1)} \tag{6.13}$$

设 α 是 $w_s^{(1)}$ 占 $w^{(1)}$ 的比例，有：

$$S_p = \frac{w_s^{(1)} + pw_{para}^{(1)}}{w_s^{(1)} + w_{para}^{(1)}} = \alpha + p(1-\alpha) \tag{6.14}$$

加速比可以随系统规模线性增大。

从另一角度看，设 $w^{(p)} = w_s^{(p)} + w_{para}^{(p)}$，则：

$$T^{(p)} = w_s^{(p)} + \frac{w_{para}^{(p)}}{p} + O(p, w^{(p)}) \tag{6.15}$$

设处理机数从 p 增加到 p' 时，把 $w^{(p)}$ 增加到 $w^{(p')}$，设此时

$$w^{(p')} = w_s^{(p')} + w_{para}^{(p')} \tag{6.16}$$

则：

$$T^{(p')} = O(p', w^{(p')}) + w_{para}^{(p')}/p' + w_s^{(p')} \tag{6.17}$$

由于 $w_s^{(p)} = w_s^{(p')}$，在时间相等的条件下有：

$$O(p, w^{(p)}) + w_{para}^{(p)}/p = O(p', w^{(p')}) + w_{para}^{(p')}/p \tag{6.18}$$

这时因为一般有 $O(p, w^{(p)}) < O(p', w^{(p')})$，从而 $\frac{w_{para}^{(p')}}{p'} < \frac{w_{para}}{p}$，我们可以定义机器对算法的可扩展性为在执行时间相等条件下负载分配到的平均处理器之比：

$$\mathrm{Scale}(T, w^{(p)}, w^{(p')}) = \frac{p/w^{(p)}}{p'/w^{(p')}} = \frac{w^{(p')}p}{p'w^{(p)}} \tag{6.19}$$

由(6.18)知：

$$w_{para}^{(p')} = \frac{p'}{p}w_{para}^{(p)} + p'(O(p, w^{(p)}) - O(p', w^{(p')})) \tag{6.20}$$

问题规模很大时，$w^{(p')} \approx w_{para}^{(p')}$，$w^{(p)} \approx w_{para}^{(p)}$ 从而有：

$$\begin{aligned}\mathrm{Scale}(T, w^{(p)}, w^{(p')}) &\approx \frac{p'w^{(p)} - pp'(O(p', w^{(p')}) - O(p, w^{(p)}))}{p'w^{(p)}} \\ &\approx 1 - \frac{p}{w^{(p)}}(O(p', w^{(p')}) - O(p, w^{(p)}))\end{aligned} \tag{6.21}$$

由(6.21)式可以看出，可扩因子是小于 1 的。可扩因子越接近 1，说明机器对算法的可扩性越好。该模型衡量了针对执行时间的可扩性。

3. *存储受限模型(加速比定律 3)*

假定 $g(pm)$ 是存储量增加 p 倍时的并行工作负载，且

$$g(pm) = G(p)g(m) \tag{6.22}$$

如果串行工作量不变，则有

$$S_p = \frac{w_s + G(p)g(m)}{w_s + G(p)g(m)/p} = p \cdot \frac{w_s + G(p)g(m)}{pw_s + G(p)g(m)} \tag{6.23}$$

因为 $G(p)g(m)$ 一般比 pw_s 大得多，所以 $S_p \approx p$，完全线性增长。

设单机存储量为 m，使 p 台处理机并行工作，占用 pm 的存储量，其工作负载为 $w^{(p)}$，则执行时间为

$$w_s^{(p)} + w_{\text{para}}^{(p)}/p + O(p, w^{(p)}) \tag{6.24}$$

在 p 增加到 p' 时，$w^{(p)}$ 也要相应增加到 $w^{(p')}$，使 $w^{(p')}$ 占用 $p'm$ 的存储量以充分利用存储资源，此时执行时间为：

$$w_s^{(p')} + w_{\text{para}}^{(p')}/p' + O(p, w^{(p')}) \tag{6.25}$$

这时可扩性就可定义为：

$$\text{Scale}(m, p, p') = \frac{w^{(p)}/p}{w^{(p')}/p'} = \frac{p'w^{(p)}}{pw^{(p')}} \tag{6.26}$$

反过来考虑，设当问题规模变化时，所要的内存量是 w 的函数，即

$$M = h(w) \tag{6.27}$$

为充分利用资源，在系统规模从 p 变到 p' 时，必然就是要使问题的平均内存消耗量相等，即

$$\frac{M^{(p)}}{p} = \frac{M^{(p')}}{p'} = m \tag{6.28}$$

即：

$$m = \frac{h(w)}{p} = \frac{h(w^{(p')})}{p'} \tag{6.29}$$

如果存在函数 f，使得取 $w = f(p)$ 时，上式对所有 p 与 p' 成立，则称 f 是等平均内存消耗函数，且这时

$$\text{Scale}(m, p, p') = \frac{p'f(p)}{pf(p')} \tag{6.30}$$

这个模型衡量的是问题求解时内存资源的利用情况，是针对内存资源利用的可扩性模型。

4. 等效率模型

并行算法在问题规模恒定时，并行效率 E 会随处理机数 p 的增加而下降，因为系统额外开销 $O_t(p, w)$（包括相对于串行算法的附加开销、打包与解包开销、通信开销、同步开销、Cache 一致性开销、内存冲突等）一般会随系统与问题规模的增加而增加。另一方面，系统规模不变时，增加问题规模，并行效率一般会提高。所以可望在系统规模扩大时，问题规模相应地适当增大，使得并行效率不变，这就是等效率模型。

$$E = \frac{w}{w + O_t(p, w)} \tag{6.31}$$

如果存在函数 f，使得当系统规模 p 增大时，问题规模以 $w = f(p)$ 的形式增长时，保持效率 E 恒定，则称 f 是等效率函数（Isoefficiency），且可扩性为：

$$\text{Scale}(E, p, p') = \frac{w^{(p)}/p}{w^{(p')}/p'} = \frac{p'w^{(p)}}{pw^{(p')}} = \frac{p'f(p)}{pf(p')} \tag{6.32}$$

如果不存在这样的函数，就认为该并行系统是不可扩的。

显然，如果 $f(g)$ 按 p 的线性函数增长，则可扩性好。如果 $f(p)$ 以 p 的多项式速度增

长，并行系统的可扩性较好。$f(p)$ 的增长速度越快，说明可扩性越差。Krustal 等人把以多项式增长的问题概括为 PE(Parallel Efficient)问题类。

杨晓东等指出用平均问题规模增长来描述可扩性具有局限性，具体表现在：

· 在共享存储结构上，不像在分布存储机器上那样容易得到程序和机器结构的开销的表达式，因为共享主存的机器上的计算涉及进程调度、Cache 一致性等，这远比分布存储的机器上的消息传递模型复杂，因而难以准确地用这种等效率模型评价共享存储结构并行系统的可扩性；

· 不能通过机器测试来评估并行系统的可扩性。

注意到这些局限性后，可以考虑不用平均负载增长量来度量可扩性，而是用平均延迟来度量，平均延迟定义为

$$L(w^{(p)},p)=\frac{1}{p}\Big(\sum_{i=1}^{p}(T_p-t_i+L_i)\Big) \tag{6.33}$$

其中 T_p 是并行算法执行时间，t_i 是处理机 i 上的执行时间，L_i 是第 i 台机上执行过程中的开销延迟(overhead latency)，包括执行过程中机器等待时间以及串行机上不需要的开销。

定义可扩展性为

$$\text{Scale}(E,p,p')=\frac{L(w^{(p)},p')}{L(w^{(p')},p')} \tag{6.34}$$

其中 $L(w^{(p)},p)$ 是保持效率 E 不变条件下的平均延迟。在 E 恒定的条件下，同时增大 p 和 w，则延迟变化越小，可扩展性越好。

在实际中 T_p,t_i,L_i 可以用软件或硬件与软件结合直接测出，从而计算出 $L(w^{(p)},p)$。等效率模型与直观性能可扩性相符，且抓住了基本特征——并行效率。

5. **等速度模型**(isospeed)

定义 1 设处理机总数为 p，取得的执行速度为 v_p，则平均单元速度为 v_p/p，简称平均速度，记为 v。

定义 2 如果系统规模增长时，能找到一种规律，当问题按这一规律增大时，保持平均速度不变，则称该并行系统可扩。

定义 3 如把系统规模从 p 扩大到 p' 时，把问题规模从 $w^{(p)}$ 扩大到 $w^{(p')}$，恰使得平均速度 v 不变，则 $\text{Scale}(v_p,p,p')=(w^{(p)}/p)/(w^{(p')}/p')=p'w^{(p)}/pw^{(p')}$。在等平均速度条件下，$\frac{w^{(p)}}{pT_p}=\frac{w^{(p')}}{p'T_{p'}}\Rightarrow\frac{p'w^{(p)}}{pw^{(p')}}=\frac{T_p}{T_{p'}}$，从而 $\text{Scale}(v,p,p')=\frac{T_p}{T_{p'}}$。

记 Scale$(1,p)$ 为 Scale(p)，则 $\text{Scale}(p)=\frac{T_1}{T_p}$(在等平均速度条件下的运行时间)。

定义 4 称并行系统的渐近平均单元速度 r_∞ 为当问题规模趋于 ∞ 时在单机上执行该算法取得的速度，称 $r_{1/2}=\frac{1}{2}r_\infty$ 为初始平均速度。

定理 1 如果一个算法在系统上负载平衡，通信代价与问题规模无关，单机上的速度随工作负载增加变化很小，则该并行系统可扩，且可扩性为：

$$\text{Scale}(p,p')=\frac{T_c(p)}{O(p)}\Big[1-\frac{\mu(w^{(p)}/p}{\mu(w^{(p')}/p')}\Big]+\frac{O(p)}{O(p')} \tag{6.35}$$

其中 $T_p = T_c(p) + O(p)$ 是并行执行时间，$T_c(p)$、$O(p)$ 分别是计算时间和开销时间，$\mu(w^{(p)})$ 是单机上工作负载为 $w^{(p)}$ 时的执行速度。

证明 $$\frac{w^{(p)}}{pT_p} = \frac{w^{(p')}}{p'T_{p'}} \Rightarrow \frac{w^{(p)}}{p(T_c(p)+O(p))} = \frac{w^{(p')}}{p'(T_c(p')+O(p'))} \tag{6.36}$$

$$\Rightarrow \frac{w^{(p)}}{p\left(\dfrac{w^{(p)}}{p\mu(w^{(p)}/p)} + O(p)\right)} = \frac{w^{(p')}}{p'\left(\dfrac{w^{(p')}}{p'\mu(w^{(p')}/p')} + O(p')\right)} \tag{6.37}$$

$$\Rightarrow \frac{1}{\mu(w^{(p)}/p)} - \frac{1}{\mu(w^{(p')}/p')} + \frac{pO(p)}{w^{(p)}} = \frac{p'O(p')}{w^{(p')}} \tag{6.38}$$

因为单机上的速度随工作负载增加时变化很小，所以，

$$w^{(p')} = p'O(p')/[1/\mu(w^{(p)}/p) - 1/\mu(w^{(p')}/p) + pO(p)/w^{(p)}] \tag{6.39}$$

把 $w^{(p')}$ 代入 $\text{Scale}(p,p') = \dfrac{p'w^{(p)}}{pw^{(p')}}$ 得到结论。

定理 2 如算法在并行机上负载平衡，而且

$$w^{(p')} = p'O(p',w^{(p')})/[1/\mu(w^{(p)}/p) - 1/\mu(w^{(p')}/p') + pO(p,w^{(p)})/w^{(p)}] \tag{6.40}$$

对任何 $w^{(p)}$ 都有正的解 $w^{(p')}$，则定理 1 的结论成立。

推论 在定理 1 或定理 2 的条件下，如系统规模趋于 ∞ 时通信开销趋于 ∞，则对任何给定的 p，有：

$$\lim_{p'\to\infty} \text{Scale}(p,p') = 0 \tag{6.41}$$

6. 开销函数模型

1989 年，Zorbas 等人提出用开销函数度量可扩性。设对含串行部分 $w_s^{(1)}$ 和并行部分 $w_{para}^{(1)}$ 的问题，算法在单机上的执行时间为 T_1，则 $T_1 = w_s^{(1)} + w_{para}^{(1)}$。该问题在 p 台处理机的理想机器上执行时间为 $T'_p = w_s^{(1)} + w_{para}^{(1)}/p$。设该问题在某含 p 台处理机的机器上执行时间为 T_p，满足 $T_p \leqslant T'_p f_o(p)$，则称满足该式的最小 $f(p)$ 为开销函数。

对任何并行系统，如果 $f_o(p)$ 是常数，则认为该系统是可扩的。如果 $f_o(p)$ 为 p 的一个增函数，则其增长率决定了并行系统的不可扩程度，增长越快，越不可扩。

$f_o(p)$ 仅反映了通信等的开销，而未反映串行部分的计算时间；而且 $f_o(p)$ 只能说明不可扩程度，不能说明并行系统的可扩程度。

7. 事件(events)函数模型

Marinescu 和 Rice 认为，单一依赖于并行软件的参数不足以分析并行系统，因为硬件可能会对性能产生很大影响。事件函数模型基于 MIMD 机器上，设有 p 个控制进程分担计算，把事件总数 g 看成 p 的函数 $g(p)$（注：一个事件即一个通信动作或一个同步动作）。在一个给定时候，一个控制进程要么正在执行计算，要么在通信或被阻塞。因而并行计算的加速比可以认为是作有效计算的控制进程的平均数量。对于给定规模的问题，如果 $g(p)$ 于 p 线性增长，则渐近加速比（处理机数 $p\to\infty$ 时的极限加速比）有一个常数上界。如果 $g(p)$ 是 p 的二次或高次多项式，则最优加速比在某个 p_{opt} 达到，且当 $p\to\infty$ 时渐近加速比为 0。

对于规模给定的问题，p_{opt} 依赖于 $g(p)$，从而 $g(p)$ 是并行计算的一个特征参数，可以

用 p_{opt} 作为可扩性的度量。p_{opt} 越大，并行系统越可扩。

8. 相对性能模型

Van Catledge 认为应以某机器为参考机去衡量其他机器的性能。并行系统的相对性能定义为应用程序在该机上取得的性能与在参考机上取得的性能的比值。等性能条件定义为在并行系统中执行应用程序以取得与参考机相同性能的机器数。

在并行系统上，$T_p = T_I(G(k)s + F(k)(1-s)/p)I$，其中 $sG(k)I$ 和 $(1-s)F(k)I$ 分别是 kI 在机器上只能串行执行的部分和可并行执行部分。这里 I 是原问题的指令总数，kI 是问题增大后的指令总数。在参考机上，$T_{p'} = T'_1(G(k)s + F(k)(1-s)/p')I$，其中 T_I、T'_I 分别是被测机与参考机的平均指令周期。满足相对性能为 1(即 $T_p = T_{p'}$ 或其他性能指标)时的 p 值可作为可扩性度量。

Nussbaum 把一个算法对一个结构的可扩性定义为在给定机器上的渐近加速比与在 EREW PRAM 上的渐近加速比的比值。这个模型可以刻画通信开销，但不能用来比较运行同一问题的两个实际的算法—结构系统，也就不可能用来有效地刻画可扩性或不可扩性，只能用来比较运行给定并行算法时的并行机结构。

Vitter 以 PRAM 计算模型为基础，定义了在其上一类有效并行的算法即 pc* 类。定义为：如果一个多项式时间计算问题(class P)在 PRAM 上有一个并行算法，使达到最差性能时的处理器数最多是问题规模的多项式函数，则称该问题属于 pc* 类。显然，认为能利用更多机器的算法可扩性更好。

9. 性能价格比模型

给定开支 c，希望购买的机器性能越高越好。以性能价格比为基础的可扩性模型刻画在给定开支的条件下，购买的机器对问题性能可扩。Kumar 认为，因为机器的价格费用与处理机速度、通信带宽、处理机数、通信信道数等许多因素有关，因而是一个复杂问题。

设让问题规模 w 和机器规模 p 同时扩展时满足 $w = f(p)$，恰保持性能价格比不变，则可称该并行系统可扩，且可扩性定义为：

$$\text{Scale}(pf/c, p, p') = \frac{w^{(p)}/p}{w^{(p')}/p'} = \frac{p'w^{(p)}}{pw^{(p')}} \tag{6.42}$$

(pf 是性能，c 是价格)。显然，$f(p)$ 越小可扩性越好，$f(p)$ 越大可扩性越差。

10. 等平均开销函数模型

1996 年，吴幸福等人得出等平均开销函数，设开销为 $O_t(p,w) = pT_p - T_1$，则平均开销 $O(p,w) = \frac{1}{p}O_t = T_p - T_1/p = T_p(1-E_p)$。可认为在给定系统上，若 p 增加，且存在函数 f，当 w 随 p 以 $f(p)$ 增加时，恰保持算法的平均开销恒定，则称并行系统是可扩的，且 f 被称为等平均开销函数。

f 的变化率越快，说明保持平均开销所要增加的问题规模越大，从而可扩性越好，反之，f 变化率越慢，认为可扩性会越差。

五、几种新的可扩性分析模型

1. 等平均 I/O 模型

若要并行处理的问题的 I/O 频繁，我们就应该衡量算法中要求的 I/O 能力与机器

I/O能力的匹配程度，因为在这种情况下一般要保持平均 I/O 能力不变(设每台机器均有I/O能力，且相等，即各机完全同构)，问题规模变化率会大于机器规模的变化率。

如果存在函数 f，当 w 以 $f(p)$ 增长时，保持 I/O 平均需求不变，则称 f 是等平均I/O函数，且可扩性定义为：

$$\text{Scale}(p,p') = \frac{p'w^{(p)}}{pw^{(p')}} \tag{6.43}$$

f 变化越大，算法对机器可扩性越差，说明对 I/O 资源利用越差，同时算法性能受 I/O影响较小。

2. 等系统利用率模型

当在机器上仅运行一个问题的程序时，重点考虑的是加速比等性能指标，但有多个程序要求分配处理机让其运行时，一种分配方案是尽量把多一点的处理机分配给利用率高的程序。这时等系统利用率模型将具有十分重要的作用。

设系统规模从 p 增大到 p' 时，为保持等系统利用率 u，则要增大问题规模。如果存在函数 f，使 w 以 $f(p)$ 增长时，u 保持不变，则称 f 是等利用率函数，同时称并行系统可扩，可扩性为：

$$\text{Scale}(p,p') = \frac{p'w^{(p)}}{pw^{(p')}} = \frac{p'f(p)}{pf(p')} \tag{6.44}$$

3. 等平均通信模型

对于基于消息传递的系统，通信问题是至关重要的问题，为此有必要关注算法要求的通信量与机器能提供的通信能力之间的匹配程度。

设系统规模从 p 增大到 p' 时，问题规模以 $w=f(p)$ 增大，恰好保持算法中的平均通信不变，则称 f 是等平均通信量函数，且称并行系统是可扩的，可扩性为：

$$\text{Scale}(p,p') = \frac{p'f(p)}{pf(p')} \tag{6.45}$$

可扩性越好，即 $f(p)$ 变化率越小，说明算法对通信资源利用率越高，同时说明系统性能受通信影响越大，提高性能很大程度上受到算法中通信需求的限制。

六、可扩性模型的一般抽象与恰当提法

1. 可扩性的恰当提法

如果要主要衡量某个参数(非 p 和非 w)，可以在保持这个参数不变的条件下，同时增大 p 和 w。如果 w 的增长率大于 p 的增长率，我们应该用 $\text{Scale}(p,p') = \frac{w^{(p)}/p}{w^{(p')}/p'} = \frac{p'w^{(p)}}{pw^{(p')}}$ 来衡量并行系统的可扩性，因为对该参数的影响主要来自 w，所以实际反映了算法对机器规模的适应程度，即表现为算法对机器的可扩性问题。

当系统规模变化快于问题规模变化时，我们可以用 $\text{Scale}(w^{(p)},w^{(p')}) = \frac{p/w^{(p)}}{p'/w^{(p')}} = \frac{pw^{(p')}}{p'w^{(p)}}$ 来衡量并行系统的可扩性，实际反映了对问题规模的增大，机器能否适应规模的变化，即表现为机器对算法的可扩性问题。

在特殊情况下，如要保持某参数不变只需使问题规模和系统规模以同样的速率增长，

则称之为完全可扩的并行系统，这时机器和问题达到了完全和谐的地步。

2. 可扩性模型的一般抽象定义

要衡量某参数，如资源利用率，可以让该参数不变，使 p 和 w 同时增长。如果 p 和 w 增长时满足关系 $w = f(p)$ 且 w 的增长率快于 p 的增长率，我们称算法对机器是可扩的，称 f 是等参数函数，且可扩性为：$\text{Scale}(p,p') = \dfrac{w^{(p)}/p}{w^{(p')}/p'} = \dfrac{p'f(p)}{pf(p')}$。

如 p 和 w 的增长满足关系 $p = g(w)$ 且 p 的增长率大于 w 的增长率，我们称机器对算法是可扩的，称 g 是等参数函数，并且可扩性为：

$$\text{Scale}(w^{(p)},w^{(p')}) = \frac{p/w^{(p)}}{p'w^{(p')}} = \frac{w^{(p')}g(w^{(p)})}{w^{(p)}g(w^{(p')})} \tag{6.46}$$

3. 可扩性模型的基本类型

(1) 性能分析型

建立在等效率、等速度、等运行时间等基础上的模型反映的是性能可扩程度，建立在开销函数等基础上的模型，反映的是系统性能的不可扩程度。

可扩程度与不可扩程度是互逆的，对同一性能指标而言如果可扩程度好，则不可扩程度就不好。

(2) 资源分析型

如内存受限模型，这种模型反映了在系统规模增大时，问题求解时对资源的利用程度，或是问题规模增大时，资源会不会成为求解时的瓶颈，以及成为瓶颈的程度。

七、并行系统可扩性模型在工作站集群(COW)上的推广

前面讨论的可扩性模型都是以完全同构的机器为基础的，难以有效地用来分析 COW 系统，主要是由于 COW 中工作站的计算能力和通信能力不尽相同，所以用平均负载的变化率不足以充分分析 COW 系统。

事实上，前面的 p 也就代表了系统的总的计算能力，前面的计算模型可以看成是在保持某参数不变情况下，问题要求的计算能力(w) 与机器的实际计算能力(用 p 刻画) 之间的匹配程度，所以推广之，可以利用计算能力需求与具体的计算能力的比值的变化率来刻画 COW 系统的可扩性。

COW 系统的计算能力可以用各工作站的计算能力之和来评估，这样就可以把用于同构系统的可扩性模型应用到 COW 系统中来。

第七章　并行编程与示范实例

本章将结合向量计算机、共享主存向量多处理机和分布存储计算机的特点，介绍与向量并行计算、共享主存向量多处理机多任务并行计算和分布式并行计算有关的基本概念以及缩短程序运行的墙上时间，提高加速比的措施。

7.1　单机向量并行编程

一、向量计算机的并行处理特性

向量处理是向量计算机最有效的特性之一，它们通常拥有以下一些支持并行处理的特性。

1. *采用流水线加工方式进行向量和标量运算*

重迭技术是提高计算机运算性能的一种有效方法，它通过将运算部件分为若干流水线站来实现。对于标量指令，第一条指令的一对操作数先进入第一站执行，当它进入第二站执行时，第二条指令的一对操作数即可进入第一站，依此类推。这样同一种操作码的若干条相继指令的运算可以同时在一个运算部件的不同站执行，它实际上起到了并行运算的作用。对于向量运算，一对对操作数一拍一拍地送入流水线，经若干拍起步时间后，运算结果相继从流水线流出。向量寄存器与存储器的数据处理也是向量操作。一个典型的向量操作序列是向量取、向量运算、向量存。

2. *功能部件独立*

向量机的 CPU 可在不同的功能部件上同时执行指令，这是因为功能部件彼此独立。功能部件的独立性使得不同操作如加、乘等可并行执行。向量编译器将分析代码以产生充分利用功能部件独立性的目标代码。

3. *向量链接技术*

在目标程序中，前一指令的结果数正好是后续指令的源操作数的现象大量存在，也是不可避免的，这种操作数据相关已成为采用多功能部件提高并行计算性能的一大障碍。为提高向量功能部件的使用效率，根据向量运算的特点，不同的向量功能部件可以采用“链接”技术予以链接，即对于遇到向量运算操作数相关的相继两条指令，只要其他运算条件已具备，就可以将这两条指令所对应的两条流水线链接在一起，前一条流水线流出的结果直接作为后一条流水线的输入操作数，从而构成了相继两条向量运算的复合流水线。

4. *短向量与长向量操作*

一个向量寄存器可存放 L 个字的数据（L 称为机器的向量长度，YH-1 的 L 为 128，YH-2 的 L 为 64），一个向量的实际长度等于或小于 L 的操作称为短向量操作，实际向量长度超过 L 的称为长向量操作。在长向量操作中，向量寄存器将分段取数据，遇到长向量，向量编译器将自动产生长向量操作指令序列。

二、编译器的优化特性

向量计算机的向量编译器具有丰富的优化特性，可自动优化 Fortran 源程序。例如对于可向量化的源代码，向量编译器将尽可能自动使其向量化，以产生较好的目标代码。如果用户能给予协助，改写程序适应编译器的优化特性，向量编译器的自动化性能则会发挥得更好。向量编译器除具有自动优化特性外还提供用户可选的优化手段——编译选择(Compiler options)和编译指导命令(Compiler directives)以及一种优化工具：流追踪(Flowtrace)，现分述如下。

1. 编译选择

用户在调用向量编译器时可通过作业控制语句指定所需的编译选择。许多涉及优化功能的任选项可由用户选用和关闭。任选是针对整个程序编译过程而言的。YH 机的编译选择参数表如下：

(1) 算术优化

INT=24,64

银河 Fortran 编译器(以下简称 YFT)对 INTEGER 定义的实体一般产生位数为 64 的整数，然而多数程序无需这样多的整数位存放整数。YFT 允许指定整数变量的长度为 24 位，这样一般能使运算加快，但也有例外，如：

```
      DO  1  I=1,N
         Y=Y+expr
1     CONTINUE
```

如果 Y 是 24 位整数变量，则 YFT 对该循环产生标量代码，而当 Y 为 64 位整数变量时，产生向量代码，用 64 位整型运算比 24 位运算反而快。

ON=P,OFF=P

在 YFT 中，双精度运算较之单精度运算慢得多。若使 YFT 将双精度变量和常数看做单精度(REAL 类型)则能大幅度地改善程序性能。仅当精度要求超过 48 位尾数(即大于 15 位十进制小数)时，才使用双精度。

(2) 循环优化

ON=V,OFF=V

决定是否向量化最内层循环。

ONROLL=n

展开循环迭代次数不大于 n 的最内层循环，缺省时 n=3，这样做可能使得紧贴最内层循环的外层循环能向量化。

OPT=BL,NOBL

决定对标量 DO 循环是否采用操作数预取，即在循环每次迭代之后，总是将下次迭代要使用的操作数预取到工作寄存器中。这样做有 20%的加速，但有时可能产生下标越界。

OPT=CVL,NOCVL

CVL 对条件向量循环根据测试情况决定让该循环向量执行还是标量执行，而 NOCVL 对条件向量循环作标量处理。

OPT=ZEROINC,NOZEROINC

一般地,DO 循环内的常增变量(CVL)在每次循环迭代时总是有一个非零增量,但也有可能增量为零,如

```
      DO  1  I=1,N
        J=J+K
        A(J)=A(J)+B(J)
      CONTINUE
```

当 K≠0 时,循环是求一个向量和,而当 K=0 时,则是对固定的 A(J)求递归和,当 OPT=NOZEROINC 时,YFT 对 DO 1 的向量化会产生错误的结果。可用 OPT=ZEROINC来禁止对包含 CIV=CIV+VAR 语句的循环向量化。

OPT=INMOV,NOINVMOV

决定是否将不变量从循环中移出。

(3) 其他

ON=F,OFF=F

决定是否使用流追踪技术。

OPT=NOIFCON,PARTIALIFCON,FULLIFCON

YFT 对形式为 IF(Condition) VAR=exp 的条件替换语句,用条件向量归并语句 Var = CVMGX(exp,Var,Condition)来替换,结果可导致这类 IF 语句的向量执行。当 OPT=FULLIFCON 时,YFT 将尽可能地用条件归并语句取代 IF 语句。当选择 PARTIALIFCON时,若 Var=exp 中的 exp 包含除法或外部函数引用,YFT 不进行代换工作。但当选择 NOIFCON,则根本不进行代换。

OPT= NODOREP,SAFEDOREP,FULLDOREP

若有可能,YFT 就用对库子程序的调用取代单行的 DO 循环。OPT=SAFEDOREP 与 OPT=FULLDOREP 的区别仅在于:若选择 SAFEDOREP,则当 DO 循环包含潜在的相关冲突或含等价变量时,不进行替换;而选择 FULLDOREP 时,不检查这些条件。

2. *编译指导命令*

编译指导命令是包含在源程序中用于指导向量编译器的命令行。许多涉及优化功能的编译命令可由用户在程序的编译过程中启用和关闭。编译指导命令仅影响到代码的一部分,通常是一个 DO 循环或优化块。例如,在 YH Fortran 编译器的一个编译指导命令行中,1 至 5 列是 CDIR $,第 6 列必须是空白,第 7 至 72 列包含零个或多个由逗号分开的编译指导命令并允许包含空白符,第 1 列的 C 保持了程序的可移植性。YH 机的编译指导命令表如下:

(1) 算术优化

INT24,INT64[LIST]

意义同编译选择时一样,LIST 中列入的是整型变量名。

(2) 循环优化

VECTOR,NOVECTOR[=n],NEXTSCALAR

用 NOVECTOR[=n]禁止对迭代次数小于或等于 n 的循环向量化,0≤n≤128,这里所指迭代次数是编译时可决定的,否则,这条命令不起作用。使用不带 n 值的 NOVEC-

TOR 命令将禁止其后的所有 DO 循环向量化，直到遇到 VECTOR 命令。VECTOR 命令使 NOVECTOR 命令失效，但它不能使编译选择 OFF=V 失效。NEXTSCALAR 指示编译程序对该命令后的第一个 DO 循环产生标量代码。

NORECURRENCE[=n]

当一个标量出现在赋值语句的等号两端时就定义了一个递归，含递归的循环称为递归循环。然而，当循环迭代次数较小时(小于 5)，递归循环的向量代码没有标量代码有效。NORECURRENCE[=n]指定一个界限值，当递归循环的迭代次数编译时已知且不大于 n 时，编译对该循环产生标量代码。若不指定 n，编译对递归循环产生标量代码。这条命令对程序段的其余部分有效，除非遇到了另一条 NORECURRENCE 命令指定了不同的 n 值。当 n=0 时，这条命令不起作用。

SHORTLOOP

指明下一个循环的迭代次数小于计算机的向量长度，不让机器判断是否要分段取数。

IVDEP

忽略向量相关性，强行向量化。

IVDMO

忽略向量相关性和存储重叠，强行向量化。

VFUNCTION

一般地，内层循环中出现外部函数调用将禁止向量化。为改善程序性能，可用 YH 汇编语言 YHAL 写出该函数的可向量化版本，并且在调用该函数的程序段中写上 VFUNCTION 命令来说明此函数是可向量化的。

UNROLL=n,ROLL,UNROLL

ROLL 和 UNROLL 在程序段中用来关闭和打开循环展开。而 UNROLL=n 命令使 YFT 对迭代次数不大于 n 的循环进行展开。

BL,NOBL,CVL,NOCVL

意义同编译任选时一样。

NODOREP,RESUMEDOREP

若有可能，YFT 就用对库子程序的调用取代单行的 DO 循环。RESUMEDOREP 与 NODOREP 在程序中分别用于恢复与关闭这种代换功能。

(3) 其他

FLOW,NOFLOW

打开和关闭追踪。

NOIFCON,RESUMEIFCON

YFT 对形式为 IF(Condition) Var=exp 的条件替换语句用条件向量归并语句 Var=CVMGX(exp,Var,Condition)来替换。NOIFCON 与 RESUMEIFCON 分别实现在程序中关闭与恢复 YFT 的替换功能。

ALLGN

使下一个标号引用，下一个 DO 循环的第一条指令或下一个入口点对准在指缓站的边界上。ALLGN 命令必须紧贴在要对准的语句之前。ALLGN 命令的用途在于它使循环

和短小的子程序一次进入指令缓冲区，从而避免了频繁的指令装入。

3. 优化工具

系统提供优化工具流追踪可帮助用户和向量编译器对程序进行优化。流追踪提供FORTRAN程序中各程序段的运行性能分析，告知用户哪些程序段在整个程序运行中最为耗时，因而可列为重点优化对象。此外，流追踪还能产生整个程序的动态调用树。

三、充分向量化提高加速比

将一个串行计算程序不作任何修改而仅仅打开自动向量化开关，得到的向量加速比一般比较低。为了提高加速比，必须根据向量机的并行处理特性及其向量编译器的优化特性对程序进行重构。因为向量化是针对最内层循环而言的，所以我们把重构程序的重点放在最内层循环上。对于程序编写者而言，他应该知道程序中哪些内层循环是可以向量化的，哪些内层循环是绝对不能向量化的，哪些内层循环是看上去不能向量化，但经过程序重构后能向量化的。在程序重构的过程中，我们主要采用了如下的一些措施尽可能多地开发最内层循环使之向量化，提高加速比。

1. 将编译任选与编译指导命令有机地结合起来使用

编译任选作用于全程序，而编译指导命令仅作用于某一个程序段，二者可以有机地结合起来使用。向量编译器总是在最普遍、最一般的情况下判断一个内层循环能否向量化，而对于程序编写者来讲总是可以在优先排除许多不需要考虑的可能性的情况下进行判断，这比让编译器判断省力得多。对于那些难于让编译器判断的最内层循环就能否向量化，以什么方式向量化尽可能明白地写上相应的编译指导命令，如VECTOR、NOVECTOR、IVDMOIVDEP和NEXTSCALAR等，这样做不容易错过可向量化的最内层循环，获得较高的加速比。

2. 循环交换与短向量操作

在一般的计算流体力学程序中，通常流向方向的网格最多，且按一般的编程习惯将它放在多层循环的最外层。以两层循环为例，具有如下结构：

```
   DO  20  I=IS,IE
   DO  20  J=JS,JE
       循环体
   CONTINUE
```

将内层循环与外层循环交换，变成：

```
   DO  20  J=JS,JE
   DO  20  I=IS,IE
       循环体
20 CONTINUE
```

内层循环的循环长度增大了，对大部分情况而言，运算结果不变，如内层循环能向量化，就能缩短程序运行的CPU时间，提高加速比。如内层循环的循环长度小于机器的向量长度，在每一个内层循环之前写上相应的短向量操作编译指导命令CDIR $ SHORTLOOP，告诉机器无须按机器的向量长度分段取数，也能带来一定的加速。

3. 内层循环展开

对于内层循环长度小(10 以内)且不宜交换内外循环的,可将内层循环展开。编译选择 UNROLL=n 可指示编译器自动展开循环长度不大于 n 的内层循环,但这种展开只对循环体中以内层循环变量为下标的数组元素位于表达式左侧的情况进行,对于其他情形必须采用手工展开的方式进行。

4. 循环替代与循环分割

内层循环如果含有 CALL 语句或 GOTO 语句是绝对不能向量化的,CALL 语句容易替代;而 GOTO 语句总是与 IF 语句配合在一起使用的,去掉 GOTO 语句的基本方法是,看这种 IF 与 GOTO 搭配起来的结构能否用不含 IF 和 GOTO 的语句替代,或者用形式为 IF(Condition)Var=exp 的条件替换语句代替。如果其他条件满足,对于前者就能直接向量化,对于后者,可用条件向量归并语句 Var=CVMGX(exp,Var,Condition)代替进行向量化。例如,在一个二维与轴对称通用的标量流体力学计算程序中,最为耗时的核心程序的内层循环中有这样的 IF 与 GOTO 搭配使用的程序段。

```
      IF(KON,EQ.0)  GOTO  85
      YY=R(I,J)
      GOTO  90
85    YY=1.0
90    CONTINUE
```

含有 GOTO 语句,不能向量化。由于 KON 不取 0 就取 1(取 0 时为二维流动,取 1 时为轴对称流动),原程序段可用语句

```
YY=FLOAT(KON)*(R(I,J)-1.0)+1.0
```

等价替代。又如,TVD 格式、NND 格式中常用到的 Minmod 函数:$Z=\mathrm{Minmod}(x,y)=(\mathrm{sign}(x)+\mathrm{sign}(y))\min(|x|,|y|)/2$。原程序段为:

```
IF(X*Y.GT.0)THEN
  Z=AMIN1(ABS(X),ABS(Y))
  IF(X.LT.0) Z=-Z
ELSE
  Z=0.0
ENDIF
```

可改写为:

```
A=AMAX1(AMAX1(X,Y),0.0)
B=AMIN1(AMIN1(X,Y),0.0)
Z=X+Y-A-B
```

又如常用的限制数据界限(A≤X≤B)的程序段,原程序段为:

```
IF(X.GT.B)  THEN
    X=B
ELSEIF(X.LT.A) THEN
    X=A
```

ENDIF

可改写为：

```
C=(A+B)/2
X=C+(X-C)/ABS(X-C)*AMIN1((B-A)/2,ABS(X-C))
```

替代后包含上述程序段的内层循环能高效向量并行。

如果内层循环的某些程序段是绝对没法向量化的，而它又仅占内层循环循环体的一小部分，此时可以采用所谓循环分割技术，即将循环体一分为二或一分为三，让该程序段作标量处理而让循环体的其他部分作向量处理。根据我们的经验，此时一般应开设少量的临时数组存放临时变量，以保持分割后的循环与循环之间的联系。例如，我们对根据 CFL 条件求Δt 的循环和包含用罗必塔(L'Hospital)法则处理对称线上的奇性的内层循环都采用了这种方法，使进行标量处理的部分达到最小。

5. *大基本块操作*

编译选择 MAXBLOCK(最大基本块)有缺省值，对 YH-2 而言，其值为 2560，循环体超过最大基本块的内层循环不能向量化。由于流体力学计算公式复杂，内层循环内容很多，内层循环体一般都很大，须在编译选择中人为设定 MAXBLOCK 的值。例如，在用 NND 格式计算化学非平衡流时，某一内层循环体长达 300 行，我们给定 MAXBLOCK 的值为 15 000，向量化效果很好。

通过采用这些措施，程序的向量加速比会大大提高，由于向量并行计算带来的程序 CPU 运行时间的缩短，故可以达到充分利用系统资源降低计算成本的目的。

7.2　共享主存向量多处理机多任务并行编程

仅具有单个 CPU 的向量机其速度和容量远远满足不了大型科学计算问题的需要，一个自然的解决方案是，让那些相对来说速度快、容量大的向量机(曾有人将其形象地称作“大象”)以某种方式连接起来形成向量多处理机系统，用某种机制使它们协同起来完成一个整体的任务，只要组织分工合理，一般来说，总是能缩短程序运行的时间(是墙上时间而不是 CPU 时间)。就 CFD 问题而言，当采用计算机辅助设计时，它能缩短型号研制周期，使大型的航天计划早日付诸实现。

一、基本术语

1. *任务(Task)*

在多处理机多任务并行数值计算中，任务是一个具有子程序形式(即用子程序来描述)并且可以调度的计算单位，它的指令是顺序执行的。任务可以创建其他任务，创建者与被创建者可以同时运行。任务执行的代码是子程序，调用一个子程序和创建一个任务来执行这个子程序，完成相同的工作，区别在于子程序调用使得该子程序的工作立即被执行，且调用者等待；而创建任务，是由系统独立地调度任务来完成该工作，该任务可以与其他任务并行，包括与调用任务并行。

2. *多任务(Multitasking)*

多任务就是构造程序使得它有多个可以并行执行的任务，一般有宏任务(Macrotask-

ing)和微任务(Microtasking)两种方法。用户通过调用一组宏库子程序来显式地创建并行实体,组织同步通讯,进行共享数据保护,称为宏任务。宏任务的开销比较大,只有对大粒度任务效果才好,一般支持子程序级的并行。微任务是用户通过一组微任务指导来指定并行,由一个 Fortran 的预处理器来解释这些指导,改写源程序,用合适的微库子程序来代替这些指导。它的开销比较小,不仅对大粒度任务效果好,对小粒度任务效果也好,主要用于循环级的并行。微任务只提供了互斥手段,没有提供进程间的同步手段,故微任务支持的并行形式只有一种,就是无嵌套的 Cobegin-Coend 形式,不能支持 Doacross 这种并行结构。

3. 作用域(Scope)

一个变量的作用域是在程序中定义和引用该变量的区域。一个任务可访问的所有变量可分为私有(Private)变量和共享(Shared)变量两种。私有变量为一个任务定义,仅在该任务中可访问;共享变量为多个任务定义,被多个任务访问。术语私有和共享是相对子任务而言的,而术语局部(Local)和全局(Global)是相对于程序单元(如子程序)来说的,全局变量是可被多个程序单元访问的变量,局部变量是只能在一个程序单元中访问的变量。因此,在多任务时,一个变量既有相对于任务而言的作用域,又有相对于程序单元而言的作用域。比如,一个变量可能对一些子程序是全局的,但私有于一个任务,这样的变量在 YH-2 中可用 TASK COMMON 块来保护和传递数据,TASK COMMON 引入了一种新的变量作用域——全局于子程序而私有于任务。

4. 临界区(Critical Region)

临界区是访问共享资源的代码段,共享资源可以是中央存储器,输入输出(I/O)文件,子程序以及其他被程序中任务共享的资源。如果包含临界区的程序模块要并行的话,我们必须对临界区代码加以控制。实现这种控制的一种方法是在进入一个临界区时置一把锁,在出临界区时清这把锁,并且所有其他并行任务进入相应临界区以前都检测该锁。不能进入临界区的任务必须等待。一方面应使临界区尽可能小,但另一方面,一个有很多小临界区的程序则有很大的锁开销,故临界区的大小与锁操作的开销应综合加以考虑。

二、宏库子程序简介

用户可通过一组在语言级调用来开发并行性的子程序集——宏任务库来实现宏任务并行计算,宏任务子程序一般可以分为任务、事件、锁和会合四大类,它提供了并行实体(任务)的控制机制,同步机制和互斥机制,而 Fortran 语言中的 Common,私有子任务被多个子程序公用的 Common 块(在 YH-2 中称作 Task Common)和参数传递则提供了并行实体间的数据通讯机制。下面以 YH-2 为例介绍几个常用宏库子程序的功能及参数说明。

1. 任务类子程序

任务类子程序主要处理任务的创建和同步,以及管理与任务有关的一些信息。

TSKSTART(task-array,name[list])

创建一个任务控制数组为 task-array 的任务,该任务完成于程序 name(list)的功能,任务一经创建,即可调度运行。name 为任务开始执行的外部入口点,它必须在调用 TSKSTART 的程序单元中用 EXTERNAL 说明。任务控制数组是一个整型数组,长度可定义为 2 或 3。

TSKWAIT(task-array)

等待任务控制数组名为 task-array 的任务完成。

2. 锁类子程序

锁类子程序主要用于代码的临界区和共享数据的保护。锁是一个特殊的整型变量，只有“Locked”和“Unlocked”两种状态，只有通过锁类子程序，才能对它进行修改，锁变量必须在 COMMON 中。

LOCKASGN(name)

将一个整型变量标记为锁。对每一个锁变量，在它被其他锁类子程序使用之前，必须先调用它。LOCKASGN 给锁赋予初始值“Unlocked”。

LOCKON(name)

若锁为“Unlocked”状态，置锁为“Locked”状态，返回调用任务。若锁已是“Locked”状态，挂起调用任务直到该锁被另一个任务清为“Unlocked”状态，然后置为“Locked”状态，返回调用任务。

LOCKOFF(name)

清锁并返回调用任务，清锁使得正在等待该锁的一个任务恢复执行。

LOCKREL(name)

清除赋给锁的标识。若锁已置或一个任务正在等待该锁，报错。

3. 事件类子程序

事件类子程序主要实现任务同步。事件是一个特殊的整型变量，只有“Posted”、“Cleared”两种状态，事件变量应放在 COMMON 中。

EVASGN(name)

将一个整型变量标记为一个事件。对于每一个事件变量，在它被其他事件类子程序使用之前，必须先调用它。EVASGN 置事件的初始状态为“Cleared”。

EVPOST(name)

置事件变量为“Posted”状态，返回调用任务。置一个事件使得所有等待该事件的任务都恢复执行。

EVWAIT(name)

若指定事件已是“Posted”状态，任务继续执行，否则挂起调用任务直至指定事件被置为“Posted”状态。EVWAIT 不改变事件的状态。

EVCLEAR(name)

将一个事件变量置为“Cleared”状态，返回调用任务，结果使得以后调用 EVWAIT 任务时，又要等待该事件重新置为“Posted”状态。

EVREL(name)

清除赋给事件的标识。若有任务正在等待该事件置为“Posted”状态，则报错。

4. 会合类子程序

会合类子程序主要用来支持一种常见的同步方式：会合，即 n 个并发任务，在会合点会合后，再恢复并行执行。

障碍是一个特殊的整型变量，它对应一个同步点，在该点处，必须指定数目的任务到

达该处后，处理才可继续。障碍只有"Closed"和"Opened"两种状态。障碍变量必须在COMMON中。

BARASGN(barriername,taskcount)

将一个整型变量标记为障碍，其初始值为"Closed"。

BARSYNC(barriername)

使任务在障碍处会合。若到达该障碍的任务(包括自己)的个数小于 taskcount 时挂起调用任务直至 taskcount 个任务到达该障碍为止，一旦有 taskcount 个任务到达该障碍，所有这 taskcount 个任务均可继续执行(此时，障碍的状态为"Opened")。这 taskcount 个任务恢复执行后，障碍又随即回到"Closed"状态。

BARREL(barriername)

清除赋予该障碍的标识，若有任务在等待该障碍，报错。

三、多向量机宏任务并行计算

1. 任务划分

一般采用静态划分的方法来保证各任务负载的基本平衡，因为各任务所要完成的工作量是预先能够确定的。以二维流体力学计算为例，对于在每一个时间步都要作 1～2 次进出口边条处理的根任务 0 和子任务 n，可以少分配一排网格的计算工作，这样，负载能基本保持平衡。

2. 同步

非定常或时间相关的流体计算属初边值问题，其并行有一个突出特点，就是从创建任务开始到整个任务结束中间对应着很多同步点。各任务必须相应完成各自的工作，在同步点会合后，才能往下继续执行。在宏任务库的四大类子程序中，任务类子程序开销最大(比其他三类子程序有时要高出三个数量级)，其次是障碍类子程序，再其次是事件类和锁类。由于任务类子程序开销大，用它来实现同步将会得不偿失，一般仅用它来创建任务以及管理与任务有关的一些信息。障碍类子程序实现同步比较简单，每一个同步点只须写一个语句 CALL BARSYNC(barriername)。对于两个处理机的计算，用两个事件或两把锁就可简单地实现同步。n 个任务在 n 台处理机上的计算，用事件类和锁类子程序进行同步，逻辑上要复杂得多。经过计算可知：一共要设置 n(n－1)把锁或 n(n－1)个事件才能严格保证同步，即在任务与任务之间都要单向地设置一把锁或一个事件。当 n 很大时，用障碍类子程序实现同步比用锁或事件类实现同步简单，而且开销小，而当 n 比较小时，可用锁或事件实现同步，由于锁类子程序与事件类子程序相比有很多相似的地方，相应子程序的开销也一样，LOCKOFF 可以改变锁的状态，而 EVWAIT 不改变事件的状态，必须紧接其后加一个 EVCLEAR 才改变事件的状态。这样在每一个同步点要多写(n－1)个 CALL EVCLEAR(name)语句，比用锁类子程序实现同步开销要略为大一点。故在我们的计算中，一般采用锁类子程序实现同步。

3. 变量类型的区分

程序能运行(不出现死锁、死循环和溢出等现象)，但运行结果与单机不一致，一种错误是多任务划分逻辑上不严格，而另外一种错误是没有很好地区分变量的类型。使用宏任务的一种常见形式是将循环的一系列迭代组织成一个子程序作为任务来执行，即将单机

程序的循环变成子程序。因此在对原程序变量的使用情况进行分析时，主要相对循环来进行。这种分析必须考虑循环中访问的所有变量以及在循环外对这些变量的所有可能访问。在串行程序中，变量的访问情况有以下几种情况：

- 局限在一个循环内；
- 作用域超出了循环，但局限于包含循环的子程序；
- 通过参数表，其作用域超出了子程序；
- 通过 COMMON 块，其作用域超出了子程序。

某些数据，既要被同一个任务中的一组子程序访问，又不能被其他任务中的任意子程序访问。YH-Fortran 引入了一种新的扩充 TASKCOMMON 来为这类变量提供共享的存储单元。用户可充分利用这一扩充进行多任务化。

某些变量或数组是临时变量或数组，如果放在 COMMON 中，当变量或数组名与其他子程序的不重复时，在单机上运行不会出问题。但多任务后，它们被多个任务共享，由于多个任务都可对它们进行修改，就有可能出错。举例来说，前面已经介绍过，将用罗必塔(L'Hospital)法则处理对称轴上的奇性的几条语句从最内层循环体分割出来进行标量处理可以提高向量加速比。但为了保持原循环体变量之间的关系，保证结果正确性，必须开设一个临时数组。在单机计算中，我们将该临时数组放到了 COMMON，且为了节约存储空间，将它定义为一维数组。多任务计算时不作改动，多机运行结果与单机的就不一致，而且每次运行的结果也不相同，究其原因就是因为犯了这类错误。修改的方法有两个，一是数组不放在 COMMON 中，只用 DIMENSION 定义，二是将数组放在 COMMON 中，但开成二维数组，这样每一个任务只在它所分配的取值范围内对该临时数组进行修改。

总之，多任务一个程序比向量化一个程序在程序重构方面以及程序调试过程中要做更多的工作。但结果是令人鼓舞的，我们曾用显式差分计算二维流体动力学问题，在两处理机上的加速比达到了 1.9 左右，四处理机上加速比达到了 3.5 左右。

7.3 PVM 环境下的分布式并行编程

一、PVM 简介

PVM 是 Parallel Virtual Machine 的缩写，它能将一组使用 Microsoft Windows 以及 UNIX 或 LINUX 系统的异构网络环境下的计算机虚拟成单个的并行计算机。PVM 由美国 Oak Ridge 国家实验室研制，是美国国家基金会资助的公开软件系统。PVM 提供在虚拟机系统的各个结点(单 CPU 或单计算机)上创建任务以及支持任务间或结点间同步和通信的手段。PVM 应用程序可以用 C 或 Fortran 编写，使用 PVM 提供的消息传递(Message Passing)机制实现并行计算。通过发送和接收消息，应用程序的多个任务(或进程)可以互相协作，并行地解决一个大型科学与工程计算或大型事务处理问题。PVM 支持在应用、机器和网络三个层次上的异构：应用的异构不言而喻；机器的异构不仅表现在当两个机器有不同的整数和浮点数据表示形式时 PVM 负责所有的数据转换，而且表现在虚拟机中的结点机可以是任何类型的计算机，包括 Von Neumman 机、向量机、工作站、服务器甚至现代意义上的并行机；网络的异构表现在 PVM 允许虚拟机定义在由不同网络构成

的互连网络上。PVM具有通用性强和系统规模小的特点，既适合TCP/IP网络环境，又适合于MPP大型并行计算机系统，当前所有的并行计算机厂商都宣布他们的产品支持PVM，一些大公司还针对其产品的特点推出自己的高效PVM产品。

二、配置和启动PVM

PVM有两种启动方式：通过控制台程序pvm启动或直接运行pvmd启动。

1．通过控制台程序pvm启动

控制台程序pvm是一个单独运行的pvm任务，它允许用户交互式启动、查询、修改虚拟机，也可以交互式启动和终止pvm任务。控制台能在虚拟机的任何结点机上多次启动和退出，而不会影响pvm和任何运行中任务的状态。当用户键入pvm命令时，将出现提示符pvm>，pvm将从标准输入设备中接收命令。下面介绍pvm命令：

（1）help：可用此命令获取关于任何pvm交互命令的信息。以相应pvm命令的名字作为help的参数，将会给用户列出如何使用此命令的详细信息。

（2）version：显示pvm的版本号。

（3）setenv：显示或设置环境变量值。

（4）conf：给出虚拟机配置的列表。

（5）add：在虚拟机中加入新的结点机，语法形式为：add hostname…。

（6）delete：从虚拟机中删除结点机，语法形式为：delete hostname…。

（7）ps：列出虚拟机中pvm任务信息，语法形式为：ps[options]。

选项（options）：

- —a：列出所有结点机任务信息
- —hhost：指定结点机，host为pvmd的tid值。
- —nhost：指定结点机，host为结点机名。
- —l：长信息列表
- —x：—a中列出的任务加上控制台任务

（8）spawn：加载一个pvm任务程序，语法形式为：spawn[options]file[arg…]。选项（options）：

- —(count)：任务数，省缺为1
- —(host)：在host指定的结点机中加载，省缺为任一结点机
- —(ARCH)：在体系结构为ARCH的结点机上加载任务
- —?：Debugger方式的任务加载
- —>(file)：重定向任务输出到控制台
- —>>(file)：重定向任务输出到文件
- —@ ：显示任务工作过程中的trace信息到控制台
- —@(file)：显示任务工作过程中的trace信息到文件中

（9）kill：终止pvm任务。

（10）reset：终止除pvm以外的所有pvm任务，清除pvm的所有内部表项和消息队列，并使监控程序处于空闲状态。

（11）quit：退出控制台，此时pvmd和pvm作业继续运行。

(12) halt：终止 pvm 虚拟机。

2. 通过 pvmd 启动

在这种方式下，用户先编好一个配置文件 hostfile，在这个文件中，描述虚拟机由哪些计算机构成，然后以此 hostfile 为参数运行 pvmd，此时，pvmd 将会根据此文件自动配置虚拟机，并分别启动这些机器上的 pvmd。用户以后可视需要再使用 pvm 去查询、配置 pvm 虚拟机。命令形式为：% pvmd hostfile，之后，pvmd 提示用户输入其他机器的口令，输入完成后拍入 Ctrl _ Z+bg，pvm 即被放在后台运行；如果不涉及用户口令问题，命令形式可改为：% pvmd hostfile &，pvm 被直接放在后台运行。

三、几个常用的 pvm 库函数

1. call pvmfmytid(tid)

功能：在第一次调用时，将调用进程注册为 pvm 进程。在每一次调用时都返回调用进程的 tid。

参数：tid——返回调用进程的 tid，负值表示出错。

2. call pvmfexit(info)

功能：通知本地 pvmd，调用进程要退出 pvm。

参数：info——返回的整数状态码，负值表示出错。

3. call pvmfspawn(task，flag，where，ntask，tids，numt)

功能：创建一个新的 pvm 进程。

参数：

task——一个字符串，存放要创建 pvm 进程所执行的可执行文件名。该可执行文件应在进程要创建的计算机上存在。pvm 查找该可执行文件的省缺位置为 $HOME/pvm3/bin/ARCH

flag——创建 pvm 进程时所用的选项。在 Fortran 中可取下列值：

· PVMDEFAULT	0	pvm 任选一台计算机，在其上创建任务
· PVMHOST	1	where 参数指出在其上创建任务的特定计算机
· PVMARCH	2	where 参数指出在其上创建任务的特定机器类型
· PVMDEBUG	4	创建任务在调试器的控制下
· PVMTRACE	8	创建任务将产生 pvm 的轨迹参数

where——一个字符串，指定在什么地方创建 pvm 进程，根据 flag 值的不同，where 可以是计算机名，也可以是机器类型名。若 flag=0，就忽略 where 参数，pvm 自己选择一个最合适的计算机。

ntask——要创建的执行同一个可执行文件的 pvm 进程数。

tids——一个具有 ntask 个元素的数组，返回时，每一个数组元素中存放一个由该调用创建的 pvm 进程的 tid。若创建某个任务出错，则在其对应元素中存放相应的错误码。

numt——整数，返回时存放实际创建的任务数，负值表示系统错误。小于 ntask 的正数表示创建部分失败。在这种情况下，用户应检查 tids 数组中的错误代码。

4. call pvmfparent(tid)

功能：返回调用进程的父进程 tid。

参数:返回调用进程的父进程的 tid。如果调用进程不是用 pvmfspawn 创建的,返回 PvmNoParent。

5. call pvmfinitsend(encoding,bufid)

功能:清除省缺的发送缓冲区并指定消息的数据格式转换方法。

参数:encoding——一个整数,用来指定下一个消息的数据格式转换方法,在Fortran中可能的使用值如下:

- PVMDEFAULT　0　如虚拟机是同构的,则不进行数据格式转换,否则采用XDR(外部数据表示)来进行数据格式转换。
- PVMRAW　1　不进行数据格式转换。
- PVMINPLACE　2　数据仍存放在原处。

6. call pvmfpack(what,xp,nitem,stride,info)

功能:将用户提供的各种类型的数据组装到活跃消息缓冲区中。

参数:

what——一个整数,用来指定被组装的数据类型,其可能的取值如下:

- STRING　0
- REAL4　4
- BYTE1　1
- COMPLEX　5
- INTEGER2　2
- REAL8　6
- INTEGER4　3
- COMPLEX16　7

xp——指向一组字节的指针,这一组字节所存放的数据可以是任何数据类型,但数据提取时也必须是相应的数据类型。

nitem——要组装的数据项数目(不是字节数)。

stride——被组装数据项的间隔。

info——返回的整数状态码,负值表示出错。

7. call pvmfsend(tid,msgtag,info)

功能:非阻塞地发送在活跃发送缓冲区中的消息。

参数:

tid——消息接收者的 tid。

msgtag——指定要发送的消息标识,接收时必须使用相同的消息标识。

info——返回的整数状态码,负值表示出错。

8. call pvmfrecv(tid,msgtag,bufid)

功能:阻塞的消息接收,阻塞调用任务直至从指定发送者发来的指定消息标识的消息到达为止,并将接收的消息存放在一个新的活跃接收缓冲区中。

参数:

tid——用户指定的要接收消息的发送者 tid,−1 为通配符,匹配任何发送者。

msgtag——指定要接收消息的消息标识，msgtag 应大于 0，msgtag=－1 表示匹配任何消息标识。

bufid——返回活跃的接收缓冲区标识号，负值表示出错。

9. call pvmfmcast(ntask,tids,msgtag,info)

功能：将活跃发送缓冲区中的数据广播给一组任务。

参数：

ntask——接收数据的任务数。

tids——一个具有 ntask 个元素的整型数组，它的每一个元素存放一个接收数据的 pvm 进程的 tid。

msgtag——指定要接收消息的消息标识，msgtag 应大于 0，msgtag=－1 表示匹配任何消息的消息标识。

info——返回值，负值表示出错。

10. call pvmfunpack (what,xp,titem,stride,info)

功能：接收消息的任务对接收到的数据进行拆包。

参数：同 call pvmfpack (what,xp,titem,stride,info)。

11. call pvmfjoingroup(group,inum)

功能：将调用进程增加到指定的进程组中。

参数：

group——存放进程组组名的字符串。

inum——返回调用进程在进程组中的实例号。

12. call pvmflvgroup(group,info)

功能：调用进程脱离指定的进程组。

参数：

group——存放进程组组名的字符串。

info——一个整数，存放该例程返回的状态码，负值表示出错。

13. call pvmfbarrier(group,count,info)

功能：阻塞调用进程直至同组中的所有进程均调用该子程序为止。

参数：

group——存放组名的字符串，由组名指定的进程组必须存在，且调用进程是该组的成员。

count——一个整数，指定参与 barrier 同步的进程数。

info——返回的整数状态码，负值表示出错。

14. call pvmfbcast(group,msgtag,info)

功能：将活跃发送缓冲区中的数据广播给由 group 指定的进程组中的所有进程。

参数：

group——存放组名的字符串，由组名指定的进程组必须存在。

msgtag——用户提供的消息消息标识，为一整数，必须大于 0，用户程序利用它来区分不同类型的消息。

四、pvm 编程举例

pvm 应用程序可用 C 或 Fortran 语言编写，通过使用 pvm 库提供的消息传递机制来实现并行计算。使用 f77 编译 pvm 应用程序，一般需连接 libfpvm3. a 和 libpvm3. a。如果一个 Fortran 程序用到了动态任务组或进程组，则它应该顺序地连接 libfpvm3. a、libgpvm3. a 和 libpvm3. a。

在 pvm 系统中，并行应用程序的结构可以具有千差万别的形式，但多数并行应用程序结构很规范，两种典型的结构是：SPMD 模式——其中每个进程大致完成相同的工作；MASTER/SLAVE 模式——多个执行计算功能的 SLAVE 进程为一个(或多个，但很少这么用)MASTER 进程工作。

1. 编译连接

以 SGI 工作站机群环境下的计算为例，当源文件 master1. f、slave1. f，master2. f 与 slave2. f、spmd1. f、spmd2. f 存放在/usr/people/pvm/pvm3/wzh 中，且按标准方式运行时，其 MAKEFILE 文件形式如下：

```
# Custom section
# Set PVM_ARCH to your architecture type (SUN4, HP9K, RS6K, SGI, etc)
# if PVM_ARCH = BSD386 then set ARCHLIB = -lrpc
# if PVM_ARCH = SGI then set ARCHLIB = -lsun
# if PVM_ARCH = I860 then set ARCHLIB = -lrpc -lsocket
# if PVM_ARCH = IPSC2 then set ARCHLIB = -lrpc -lsocket
# otherwise leave ARCHLIB blank
#
# PVM_ARCH and ARCHLIB are set for you if you use 'aimk'.
#
PVM_ARCH    =   SGI
ARCHLIB     =   -lsun
# END of custom section - leave this line here
#
CC=    cc
F77    =    f77
PVMDIR  =   /usr/people/pvm/pvm3
PVMLIB  =   $(PVMDIR)/lib/$(PVM_ARCH)/libpvm3.a
SDIR    =   $(PVMDIR) /wzh
BDIR    =   $(PVMDIR)/bin
XDIR    =   $(BDIR)/$(PVM_ARCH)
CFLAGS  =   -g -I$(PVMDIR)/include
LIBS    =   $(PVMLIB) $(ARCHLIB)
FFLAGS  =   -g
FLIBS   =   $(PVMDIR)/lib/$(PVM_ARCH)/libfpvm3.a $(LIBS)
```

```
master1: $(SDIR)/master1.f
	$(F77) $(CFLAGS) -o master1 $(SDIR)/master1.f $(FLIBS)
mv master1 $(XDIR)
slave1: $(SDIR)/slave1.f
	$(F77) $(CFLAGS) -o slave1 $(SDIR)/slave1.f $(FLIBS)
	mv slave1 $(XDIR)
master2: $(SDIR)/master2.f
	$(F77) $(CFLAGS) -o master2 $(SDIR)/master2.f $(FLIBS)
	mv master2 $(XDIR)
slave2: $(SDIR)/slave2.f
	$(F77) $(CFLAGS) -o slave2 $(SDIR)/slave2.f $(FLIBS)
	mv slave2 $(XDIR)
spmd1: $(SDIR)/spmd1.f
	$(F77) $(CFLAGS) -o spmd1 $(SDIR)/spmd1.f $(FLIBS)
	mv spmd1 $(XDIR)
spmd2: $(SDIR)/spmd2.f
	$(F77) $(CFLAGS) -o spmd2 $(SDIR)/spmd2.f $(FLIBS)
mv spmd2 $(XDIR)
```

2. 用 Fortran 语言编写的 pvm 示范程序举例(MASTER/SLAVE 模式)

```
      program master1
      include '../include/fpvm3.h'
c ---------------------------------------------------
c Example fortran program illustrating the use of PVM 3.0
c ---------------------------------------------------
      integer i, info, nproc, msgtype
      integer mytid, tids(0:32)
      integer who
      double precision result(32), data(100)
      character*12 nodename
      character*8 arch
c ---------- Starting up all the tasks ----------
c     Enroll this program in PVM
      call pvmfmytid( mytid )
c     Initiate nproc instances of slave1 program
      print *,'How many slave programs (1-32)? '
      read *, nproc
c
c     If arch is set to '*' then ANY configured machine is acceptable
```

```
c       otherwise arch should be set to architecture type you wish to use.
        nodename = 'slave1'
        arch = '*'
        call pvmfspawn( nodename, PVMDEFAULT, arch, nproc, tids, info )
        do 100 i=0, nproc-1
          print *,'tid',i,tids(i)
100     continue
c ------- Begin user program --------
        n = 10
c       Initiate data array
        do 20 i=1,n
          data(i) = 1
20      continue
c       broadcast data to all node programs
        call pvmfinitsend( PVMDEFAULT, info )
        call pvmfpack( INTEGER4, nproc, 1, 1, info )
        call pvmfpack( INTEGER4, tids, nproc, 1, info )
        call pvmfpack( INTEGER4, n, 1, 1, info )
        call pvmfpack( REAL8, data, n, 1, info )
        msgtype = 1
        call pvmfmcast( nproc, tids, msgtype, info )
c       wait for results from nodes
        msgtypa = 2
        do 30 i=1,nproc
           call pvmfrecv( -1, msgtype, info )
           call pvmfunpack( INTEGER4, who, 1, 1, info )
           call pvmfunpack( REAL8, result(who+1), 1, 1, info )
           print *, 'I got',result(who+1), ' from', who
30      continue
c --------- End user program --------
c       program finished leave PVM before exiting
        call pvmfexit(info)
        stop
        end

        program slave1
        include '../include/fpvm3.h'
c ------------------------------------------------------------
```

```
c Example fortran program illustrating use of PVM 3.0
c ------------------------------------------------------
        integer info, mytid, mtid, msgtype, me
        integer tids(0:32)
        double precision result, data(100)
        double precision work
c  Enroll this program in PVM
        call pvmfmytid( mytid )
c  Get the master's task id
        call pvmfparent( mtid )
c ------- Begin user program --------
c       Receive data from host
        msgtype = 1
        call pvmfrecv( mtid, msgtype, info )
        call pvmfunpack( INTEGER4, nproc, 1, 1, info )
        call pvmfunpack( INTEGER4, tids, nproc, 1, info )
        call pvmfunpack( INTEGER4, n, 1, 1, info )
        call pvmfunpack( REAL8, data, n, 1, info )
c       Determine which slave I am (0 -- nproc-1)
        do 5 i=0, nproc-1
           if( tids(i) .eq. mytid ) me = i
5       continue
c       Do calculations with data
        result = work( me, n, data, tids, nproc )
c       Send result to host
        call pvmfinitsend( PVMDEFAULT, info )
        call pvmfpack( INTEGER4, me, 1, 1, info )
        call pvmfpack( REAL8, result, 1, 1, info )
        msgtype = 2
        call pvmfsend( mtid, msgtype, info )
c --------- End user program --------
c       Program finished. Leave PVM before exiting
        call pvmfexit(info)
        stop
        end

        double precision function work( me, n, data, tids, nproc )
        include '../include/fpvm3.h'
```

```
c       -----------------------------------------
c       Just a simple routine for illustration
c       -----------------------------------------
        double precision data( * ), sum, psum
        integer i, n, me, inum, nproc
        integer tids(0: * )
        sum = 0.0
        do 10 i=1,n
          sum = sum + me * data(i)
10      continue
c       -----------------------------------------
c       Pass partial result to neighboring node
c       to illustrate node-to-node communication
c       -----------------------------------------
        call pvmfinitsend( PVMDEFAULT, info )
        call pvmfpack( REAL8, sum, 1, 1, info)
        inum = me+1
        if( inum .eq. nproc ) inum = 0
        call pvmfsend( tids(inum), 77, info )
        call pvmfrecv( -1, 77, info )
        call pvmfunpack( REAL8, psum, 1, 1, info)
        work = sum + psum
        return
        end
```

7.4 MPI 环境下的分布式并行编程

MPI 是消息传递界面(Massage Passing Interface)的简称。为了统一互不兼容的用户界面,1992 年成立了 MPI 委员会,负责制定消息传递的新标准,

一、IN、OUT、INOUT 约定

MPI 用 IN、OUT、INOUT 来描述 MPI 接口中参数的属性,可解释为:

- 函数调用中只使用不被修改的参数,标记为 IN;
- 函数调用中可能被修改的参数,标记为 OUT;
- 函数调用中不仅使用而且可能被修改的参数,标记为 INOUT。

二、与 MPI 函数有关的几个重要术语

1. 非阻塞

如果一个函数调用在返回之前即允许用户使用调用中涉及的资源(如缓冲区),则称之为非阻塞。

2. 阻塞

函数调用的返回则意味着本次调用涉及的资源可重新使用。

3. 本地性

函数的执行完成无须与其他进程通信,独立于其他进程的执行。

4. 非本地性

函数的执行完成与其他进程的执行相关,需要进程间通信。

5. 全局性

函数的执行需要一个进程组中的所有进程都参与。

三、基本概念

1. 模糊对象

MPI 中的组、通信器等概念都是用对象来表示的,其内部数据表示对用户不可见,但向用户提供句柄(handle),用户可通过在 MPI 函数的参数中传递句柄实现对模糊对象的访问。

2. 命名常量

MPI 给某些类型参数的特殊值赋予的一个特殊的含义,如 MPI_ANY_TAG;MPI 还提供了一些预定义的对象常量,如 MPI_COMM_WORLD。

3. 消息描述

MPI 用一个三元组来描述一个消息,其中包括消息起始地址、数据类型、消息元素个数。MPI 用消息数据类型和消息元素个数来表示一个消息的长度,而不是直接用 byte 来表示。MPI 定义了一些基本的数据类型,也提供了构造用户自定义数据类型的接口。消息元素个数可以为零。

4. 消息信封

为了识别和区分消息,MPI 用消息信封来标识一个消息。消息信封实际上就是几个域的组合,其中需要用户填写的包括:信息源进程标识(或目的进程标识)、消息 tag 和通信器。

5. 通信器(commuticator)和组(group)

通信器表示一个通信域,用以保证通信安全,保证不同通信域传递的消息不相互干扰。MPI 预定义了一个通信器(MPI_COMM_WORLD),该通信器在 MPI 启动时产生,其中包含了并行作业中的所有进程。用户可创建自己的通信器,以表示特定的通信域。

通信器通常和进程组相关联。进程组作为通信器的一个属性,描述了该通信域中有哪些进程。后面将详细描述 MPI 中有关通信器和进程组的接口。

6. 进程标识

MPI 中用进程在组中的 rank 来标识该进程。合法的 rank 值为{0,…,n−1}U{MPI_ANY_SOURCE},n 是组中的进程数,MPI_ANY_SOURCE 只能在接收时使用。通信器参数即隐含了一个进程组。

7. 消息 tag

消息信封中的消息 tag 是一个整数,应用程序可以用它来区分消息。合法的 tag 值的范围为{0,…,UB}U{MPI_ANY_TAG},UB 是最大 tag 值,MPI_ANY_TAG 只能在

接收时使用。UB 的值是实现相关的，用户可通过查询 MPI_TAG_UB 属性获取 UB 值。

四、点点通信

最基本的点点通信操作就是 send 和 receive。MPI 提供了多种语义不同的点点通信操作，包括阻塞、非阻塞和不同通信模式下的通信操作。下面介绍 Fortran 程序接口。

1. 阻塞的消息发送和接收

(1) MPI_SEND——基本的阻塞消息发送(非本地操作)

· 接口描述

```
MPI_SEND(BUF,COUNT,DATATYPE,DEST,TAG,COMM,IERROR)
POINTER BUF(*)
INTEGER COUNT,DATATYPE,DEST,TAG,COMM,IERROR
```

· 参数描述

IN	BUF	发送缓冲区的起始地址
IN	COUNT	发送缓冲区中消息元素的个数(非负整数)
IN	DATATYPE	发送缓冲区中消息元素的数据类型
IN	DEST	目的进程 Rank
IN	TAG	消息标识
IN	COMM	通信器
OUT	IERROR	返回的错误码

· 返回错误描述

MPI_SUCCESS	无错，MPI 函数执行正确
MPI_ERR_COMM	非法通信器
MPI_ERR_COUNT	非法消息元素个数
MPI_ERR_TYPE	非法消息元素类型
MPI_ERR_TAG	非法消息标识
MPI_ERR_RANK	非法进程 Rank

· 语义描述

MPI_SEND 按参数中指定的消息信封将消息发往目的进程，返回后用户可自由存取或覆盖消息发送缓冲区，这时有两种可能：消息已经直接拷贝到接收方缓冲区；或者消息被放入一个临时的系统缓冲区。具体采用何种方式由系统决定。

(2) MPI_RECV——阻塞的消息接收操作(非本地操作)

· 接口描述

```
MPI_RECV(BUF,COUNT,DATATYPE,SOURCE,TAG,COMM,STATUS,IERROR)
POINTER BUF(*)
INTEGER COUNT,DATATYPE,SOURCE,TAG,COMM,STATUS(MPI_STATUS_SIZE),IERROR
```

· 参数描述

OUT	BUF	接收缓冲区的起始地址

IN	COUNT	接收缓冲区中能容纳的消息元素的最大个数
IN	DATATYPE	接收缓冲区中消息元素的数据类型
IN	SOURCE	信息源的进程 Rank
IN	TAG	消息标识
IN	COMM	通信器
OUT	STATUS	状态对象
IN	IERROR	返回的错误码

· 返回错误描述

参照 MPI _ SEND 的返回错误描述

· 语义描述

MPI _ RECV 按参数中指定的消息信封接收相应的消息，正常返回则意味着消息已拷贝进用户提供的接收缓冲区；如果用户提供的接收缓冲区长度小于接收的消息，MPI 直接返回一个 overflow 错。Status 对象由用户分配，它向用户返回一些状态信息。

2. 非阻塞的消息发送和接收

MPI 提供对非阻塞通信的支持，使用户获得性能上的好处。一个完整的 MPI 的非阻塞通信过程包括两个分离的阶段：非阻塞地提交通信请求，然后在某个时机再去阻塞地确认它的完成。

(1) MPI _ ISEND——开始提交一个非阻塞发送

· 接口描述

MPI _ ISEND (BUF, COUNT, DATATYPE, DEST, TAG, COMM, REQUEST, IERROR)

POINTER BUF(*)

INTEGER COUNT,DATATYPE,DEST,TAG,COMM,REQUEST,IERROR

· 输入参数描述

参照 MPI _ SEND 的参数描述。

· 输出参数描述

OUT　　REQUEST　　发送请求(handle)

· 返回错误描述

MPI _ ERR _ INTERN　　可能的内存申请错误

其他错误参照 MPI _ SEND 的返回错误描述。

· 语义描述

MPI _ ISEND 提交一个标准模式的发送，向用户返回一个发送请求的 handle，该 handle使用户可以查询该发送请求的状态，并且完成该发送请求。MPI _ ISEND 的返回仅意味着消息已经在发送中，此时用户不可以重用消息的发送缓冲区，直到显式地用 MPI _ TEST 或 MPI _ WAIT 阻塞地完成消息发送为止。

(2) MPI _ IRECV——开始一个非阻塞接收

· 接口描述

MPI _ IRECV (BUF, COUNT, DATATYPE, SOURCE, TAG, COMM, REQUEST,

IERROR)

POINTER BUF(*)

INTEGER COUNT,DATATYPE,SOURCE,TAG,COMM,REQUEST,IERROR

· 输入参数描述

参照 MPI _ RECV 的参数描述。

· 输出参数描述

OUT REQUEST 接收请求(handle)

· 返回错误描述

参照 MPI _ ISEND 的错误描述

· 语义描述

MPI _ IRECV 提交一个接收请求,并向用户返回一个指向该请求的 handle,该 handle 使用户可以查询该接收请求的状态,并且完成该接收请求。MPI _ IRECV 的返回仅意味着消息已经在接收中,此时用户不可以使用消息的接收缓冲区,直到显式地用 MPI _ TEST 或 MPI _ WAIT 阻塞地完成消息接收为止。

(3) MPI _ WAIT——等待 MPI 消息发送或接收操作的完成

· 接口描述

MPI _ WAIT(REQUEST,STATUS,IERROR)

INTEGER REQUEST,STATUS(MPI _ STATUS _ SIZE),IERROR

· 输入参数描述

INOUT REQUEST 提交非阻塞操作返回的发送或接收请求的 handle

· 输出参数描述

OUT STATUS 返回该请求的状态消息

· 返回错误描述

MPI _ SUCCESS MPI 操作成功地完成

MPI _ ERR _ REQUEST 非法地请求对象

· 语义描述

当 REQUEST 对象标识的非阻塞操作完成后,MPI _ WAIT 返回,并且由 MPI _ WAIT 释放 REQUEST 对象,并将其置为 MPI _ REQUEST _ NULL。MPI _ WAIT 返回时,STATUS 对象包含了被完成操作的一些信息。MPI _ WAIT 具有非本地操作的语义。

(4) MPI _ TEST——检测非阻塞操作是否完成

· 接口描述

MPI _ TEST(REQUEST,FLAG,STATUS,IERROR)

LOGICAL FLAG

INTEGER REQUEST,STATUS(MPI _ STATUS _ SIZE),IERROR

· 输入参数描述

INOUT REQUEST 提交非阻塞操作返回的发送或接收请求的 handle

· 输出参数描述

OUT FLAG 如果非阻塞操作完成,FLAG 为真

OUT　　STATUS　返回该请求的状态消息

· 返回错误描述

MPI _ SUCCESS　　　　MPI 操作成功地完成

MPI _ ERR _ REQUEST　　非法地请求对象

· 语义描述

当 REQUEST 对象标识的非阻塞操作完成后，MPI _ TEST 返回 FLAG=TRUE，并且由 MPI _ TEST 释放 REQUEST 对象；否则返回 FLAG=FALSE，这时 STATUS 无有效值。MPI _ TEST 具有本地操作的语义。

(5) MPI _ WAITALL——等待 MPI 消息发送或接收操作的完成

· 接口描述

MPI _ WAITALL (COUNT, ARRAY _ OF _ REQUESTS, ARRAY _ OF _ STATUSES, IERROR)

INTEGER COUNT, ARRAY _ OF _ REQUESTS(*), ARRAY _ OF _ STATUSES (MPI _ STATUS _ SIZE, *), IERROR

· 输入参数描述

IN　　COUNT　　　　数组长度

INOUT　　ARRAY _ OF _ REQUESTS　　非阻塞请求数组

· 输出参数描述

OUT　　ARRAY _ OF _ STATUSES 返回的状态对象数组

· 返回错误描述

MPI _ SUCCESS　　　　MPI 操作成功地完成

MPI _ ERR _ REQUEST　　非法地请求对象

· 语义描述

阻塞地等待 REQUEST 数组关联的所有非阻塞通信操作完成，有关的返回信息放在 STATUS 对象数组中，并释放所有的 REQUEST 对象。

(6) MPI _ TESTALL——检测非阻塞操作是否完成

· 接口描述

MPI _ TESTALL (COUNT, ARRAY _ OF _ REQUESTS, FLAG, ARRAY _ OF _ STATUSES, IERROR)

LOGICAL FLAG

INTEGER COUNT, ARRAY _ OF _ REQUESTS(*), ARRAY _ OF _ STATUSES (MPI _ STATUS _ SIZE, *), IERROR

· 输入参数描述

IN　　COUNT　　　　数组长度

INOUT　　ARRAY _ OF _ REQUESTS　　非阻塞请求数组

· 输出参数描述

OUT　　FLAG　　　　如果请求数组中的所有非阻塞操作均完成，则返回 TRUE

OUT　　ARRAY_OF_STATUSES　　返回的状态对象数组

· 返回错误描述

MPI_SUCCESS　　MPI 操作成功地完成

MPI_ERR_REQUEST　　非法地请求对象

· 语义描述

MPI_TESTALL 非阻塞地测试请求数组中的所有非阻塞操作，如果所有的非阻塞操作均完成，则置 FLAG 为 TRUE，并释放 REQUEST 对象，返回的 STATUS 对象放在 STATUS 数组中，分别对应一个已完成的 REQUEST；否则，置 FLAG 为 FALSE。

五、不同通信模式下消息的发送和接收

MPI 提供了几种点点通信模式供用户使用，用户可以根据系统平台和并行应用的特点来选择适当的通信模式，以获取性能上的好处。无论是阻塞还是非阻塞方式 MPI 都提供了四种通信模式：标准(Standard)模式、缓冲(Buffered)模式、同步(Synchronous)模式和就绪(Ready)模式。

1. *标准模式*

对于标准模式，由 MPI 系统决定是否缓冲将要发送的消息。如果系统缓冲区足够用，MPI 可能缓冲消息，调用直接返回，这时相应的接收操作可能还没有发生。如果系统认为缓冲区不足，MPI 就会选择不缓冲消息，阻塞到消息被发送为止。前面描述的就是标准模式的接口。建议使用标准模式。

2. *缓冲模式*

缓冲模式要求用户提供足够的缓冲区，MPI 将需要发送的消息缓冲起来，直接返回，而不依赖相应的接收操作是否发生；如果当前没有足够的缓冲区，则返回一个错误信息。

3. *同步模式*

同步模式的发送操作不依赖于相应的接收操作是否已发生，但是该发送只能等到相应的接收操作已经在接收消息后才能返回。因此同步发送的返回不仅意味着发送缓冲区可重用，也指示消息相应的接收操作已在进行。

4. *就绪模式*

只有当相应的接收已经提交后，用户才能启动就绪模式的消息发送，否则该操作的结果不可预测。就绪模式在实现时直接向接收方发消息而没有握手过程。

六、全局通信

MPI 的全局通信操作中对消息的描述和点点通信一样。MPI 的全局操作都具有阻塞的语义，没有非阻塞的语义。下面描述主要的全局通信接口。

1. MPI_BARRIER

· 接口描述

MPI_BARRIER(COMM,IERROR)

INTEGER COMM,IERROR

· 返回错误描述

MPI_SUCCESS　　MPI 调用成功完成；

MPI_ERR_COMM　　非法通信器参数，可能使用了一个不存在的通信器 handle；

· 语义描述

用以在通信器参数隐含的组内进行障碍同步，当组内所有的进程都开始调用 MPI _ BARRIER 时，MPI _ BARRIER 返回。

2. MPI _ BCAST

· 接口描述

MPI _ BCAST(BUFFER, COUNT, DATATYPE, ROOT, COMM, IERROR)

POINTER BUFFER(*)

INTEGER COUNT, DATATYPE, ROOT, COMM, IERROR

· 输入参数描述

INOUT	BUFFER	消息缓冲区首地址
IN	COUNT	消息元素的个数
IN	DATATYPE	消息数据类型
IN	ROOT	广播操作的根进程在组内的标识
IN	COMM	用以标识组的通信器

· 返回错误描述

MPI _ ERR _ ROOT　　错误的根进程标识。ROOT 必须是组内合法的标识

· 语义描述

MPI _ BCAST 将根进程在接口参数中指定的消息向组内所有的成员广播。参与广播的进程都调用 MPI _ BCAST，根进程接口参数中的缓冲区是消息发送缓冲区，其他进程的接口参数中描述的则是接收缓冲区。发送方和接收方可使用不同的数据类型，但发送的数据量和接收的数据量必须相同。

3. MPI _ GATHER

· 接口描述

MPI _ GATHER (SENDBUF, SENDCOUNT, SENDTYPE, RECVBUF, RECVCOUNT, RECVTYPE, ROOT, COMM, IERROR)

POINTER　SENDBUF(*), RECVBUF(*)

INTEGER　SENDCOUNT, SENDTYPE, RECVCOUNT, RECVTYPE, ROOT, COMM, IERROR

· 输入参数描述

IN	SENDBUF	发送缓冲区首地址
IN	SENDCOUNT	发送消息的元素个数
IN	SENDTYPE	发送消息的数据类型
OUT	RECVBUF	接收缓冲区首地址
IN	RECVCOUNT	接收消息的元素个数(在根进程中才有效)
IN	RECVTYPE	接收消息的数据类型(在根进程中才有效)
IN	ROOT	根进程(接收进程)在组内的标识
IN	COMM	用以标识组的通信器

· 输出参数描述

OUT　RECVBUF　　接收缓冲区首地址(在根进程中才有效)

・返回错误描述(略)

・语义描述

组内所有的进程(包括根进程)将其接口参数中描述的消息发往根进程,根进程接收消息,并按组内 RANK 的顺序将接收到的每个消息放入接收缓冲区。根进程的 RECVCOUNT 指的是从每个其他进程接收的消息的数量,而不是它接收的所有消息的数量。

4. MPI_SCATTER(执行与 MPI_GATHER 相反的操作)

・接口描述

MPI_SCATTER(SENDBUF, SENDCOUNT, SENDTYPE, RECVBUF, RECVCOUNT, RECVTYPE, ROOT, COMM, IERROR)

POINTER SENDBUF(*), RECVBUF(*)

INTEGER SENDCOUNT, SENDTYPE, RECVCOUNT, RECVTYPE, ROOT, COMM, IERROR

・输入参数描述

IN	SENDBUF	发送缓冲区首地址(在根进程中才有效)
IN	SENDCOUNT	发送给其他每个进程的消息的元素个数(在根进程中才有效)
IN	SENDTYPE	发送缓冲区中消息的数据类型(在根进程中才有效)
OUT	RECVBUF	接收缓冲区首地址
IN	RECVCOUNT	要接收的消息的元素个数
IN	RECVTYPE	要接收的消息的数据类型
IN	ROOT	根进程(接收进程)在组内的标识
IN	COMM	用以标识组的通信器

・输出参数描述

OUT　RECVBUF　　接收缓冲区首地址

・返回错误描述(略)

・语义描述

MPI_SCATTER 与 MPI_GATHER 的语义相反。在 MPI_SCATTER 操作中,根进程将消息发送缓冲区中的消息切成 n 块(假设组内有 n 个进程),其中第 i 块发给组内第 i 个进程(进程标识为 i),这里包括发给根进程自己的部分。

七、通信器和组的有关操作

MPI 提供了通信器和进程组的概念。MPI 程序在调用 MPI_INIT 初始化后,即存在一个预定义的通信器(MPI_COMM_WORLD)和与该通信器关联的进程组,该进程组包含参与该并行应用的所有进程。

1. 对组的操作

(1) MPI_GROUP_SIZE

・接口描述

MPI_GROUP_SIZE(GROUP,SIZE,IERROR)

INTEGER GROUP,SIZE,IERROR

· 输入输出参数描述

IN	GROUP	组的句柄
OUT	SIZE	组内的进程数

· 语义描述

该接口返回组内的进程数，如果组为 MPI_GROUP_EMPTY，则返回 SIZE=0。

(2) MPI_GROUP_RANK

· 接口描述

MPI_GROUP_RANK(GROUP,RANK,IERROR)

INTEGER GROUP,RANK,IERROR

· 输入输出参数描述

IN	GROUP	组的句柄
OUT	RANK	该进程在组内的标识(RANK)

· 语义描述

该接口返回调用进程在组内的标识(RANK)。

(3) MPI_GROUP_FREE

· 接口描述

MPI_GROUP_FREE(GROUP,IERROR)

INTEGER GROUP,IERROR

· 输入输出参数描述

IN GROUP 被释放的组的句柄

· 语义描述

释放组 GROUP，置为 MPI_GROUP_NULL，任何与其相关的通信操作都将完成。

2. 对通信器的操作

(1) MPI_COMM_SIZE

· 接口描述

MPI_COMM_SIZE(COMM,SIZE,IERROR)

INTEGER COMM,SIZE,IERROR

· 输入输出参数描述

IN	COMM	通信器
OUT	SIZE	返回通信器所关联的组的大小

· 返回错误描述

MPI_SUCCESS MPI 调用成功完成；

MPI_ERR_COMM 通信器非法；

· 语义描述

返回通信器所包含的进程个数。

(2) MPI_COMM_RANK

· 接口描述

MPI _ COMM _ RANK(COMM,RANK,IERROR)

INTEGER COMM,RANK,IERROR

· 输入输出参数描述

IN	COMM	通信器
OUT	RANK	本进程在通信器所关联的组中的标识(RANK)

· 返回错误描述(略)

· 语义描述

该接口返回本调用进程的 RANK(0,1,…,SIZE−1,其中 SIZE 为 MPI _ COMM _ SIZE 的返回值),且该 RANK 相对于通信器所关联的组。

(3) MPI _ COMM _ FREE

· 接口描述

MPI _ COMM _ FREE(COMM,IERROR)

INTEGER COMM,IERROR

· 输入输出参数描述

IN	COMM	通信器

· 语义描述

释放组一个通信 GROUP,COMM 被置为 MPI _ CMM _ NULL,同时与该通信器所关联的所有通信操作均将被完成。

八、环境管理

MPI 允许用户获取和设置某些执行环境的参数,并且参与错误处理。MPI 在初始化时将一系列属性关联到预定义的通信器 MPI _ COMM _ WORLD,用户可通过下列函数获取这些属性。

1. MPI _ GET _ PROCESSOR _ NAME

· 接口描述

MPI _ GET _ PROCESSOR _ NAME(NAME,RESULTLEN,IERROR)

CHARACTER NAME

INTEGER RESULTLEN,IERROR

· 输入输出参数描述

OUT	NAME	处理机名
OUT	RESULTLEN	字符串 NAME 的长度

· 语义描述

该接口返回调用进程所处的处理机名。

2. MPI _ WTIME

· 接口描述

DOUBLE PRECISION MPI _ WTIME( )

· 输入输出参数描述(无)

· 语义描述

以浮点数返回以秒为单位的墙上时间。

3. MPI _ INIT

· 接口描述

MPI _ INIT(IERROR)

INTEGER IERROR

· 语义描述

该接口必须在所有 MPI 其他调用之前调用，用于完成 MPI 的初始化工作。

4. MPI _ FINALIZE

· 接口描述

MPI _ FINALIZE(IERROR)

INTEGER IERROR

· 语义描述

该接口在程序的最后调用，用于中止 MPI 的运行，释放 MPI 资源。该调用之后的所有 MPI 调用将无效。

九、Fortran 接口实例

```
      PROGRAM FORTRAN _ PARALLEL
      PARAMETER(NX=81 * 322 * 4, NUMP=4)
      INCLUDE 'MPIF. H'
      REAL A(NX, (NUMP-1) * 40)
      REAL B(NX,40)
      REAL * 8 T0
      INTEGER IR(120)
      INTEGER IS(120)
      INTEGER STATUS(MPI _ STATUS _ SIZE,120)
      INTEGER STATUSA(MPI _ STATUS _ SIZE)
      CALL MPI _ INIT(IERR)
      CALL MPI _ COMM _ RANK(MPI _ COMM _ WORLD,MYTID,IERR)
      CALL MPI _ COMM _ SIZE(MPI _ COMM _ WORLD,NUMP,IEER)

      IF (MYTID. NE. 0) THEN
        DO J=1,40
        DO I=1,NX
          B(I,J)=J * 1. 0
        ENDDO
        ENDDO
      ENDIF
C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
      CALL MPI _ BARRIER(MPI _ COMM _ WORLD,IERR)
```

```
WRITE( * , * ) 'MYTID=',MYTID

MPI_D=MPI_REAL
T0=MPI_WTIME()
T1=T0
WRITE( * , * ) T0,T1,REAL(T0)

IF (MYTID.NE.0) THEN
  DO II=1,40
   K=(MYTID-1) * 40+II
   CALL MPI_SEND(B(1,II),NX,MPI_D,0,K,MPI_COMM_WORLD,
   IERR)
  ENDDO
ENDIF

IF (MYTID.EQ.0) THEN
  T0=MPI_WTIME()
  DO I=1,NUMP-1
  DO J=1,40
   K=(I-1) * 40+J
CALL MPI_IRECV(A(1,K),NX,MPI_D,I,K,MPI_COMM_WORLD,IS
 (K),IERR)
 ENDDO
 ENDDO
 CALL MPI_WAITALL((NUMP-1) * 40,IS,STATUS,IERR)
 WRITE(1, * ) STATUS
 WRITE( * , * ) MPI_WTIME()-T0
ENDIF
CALL MPI_BARRIER(MPI_COMM_WORLD,IERR)
CALL MPI_FINALIZE(IERR)
STOP
END
```

第八章　CFD 并行计算

采用有限差分、有限体积、有限元等方法对 Navier-Stokes 方程进行离散后，有的可以直接并行，但大部分将转化为所谓强数据相关问题的求解，如大型稀疏线性代数方程的求解等。如果采用显格式进行计算，可以直接进行任务划分，局部通信多、全局通信少，并行效率高。但显格式是条件稳定的，计算时间步长受限于 C. F. L. 条件，得到一个收敛解所需的迭代步数多，模拟时间长，总体计算效率不高。可以采用当地时间步长法或加焓阻尼的办法加快收敛。隐格式具有较好的稳定性，但求解复杂，涉及稀疏带状或块带状线性方程组的求解，计算分为分裂法、迭代法和直接法三大类。分裂法的基本思想是将稀疏五对角块和稀疏七对角块线性方程组的求解转化为2～3个三对角块线性方程组的求解，如交替方向隐格式 ADI 和近式因子隐格式 AF 等；或转化为除法运算，如 LU 分解等。迭代法包括 Jacobi 迭代、Gauss-Seidel 迭代、亚松弛迭代、超松弛迭代以及点松弛、线松弛和面松弛方法等。直接法就是用消去和回代的方法理论上无误差地直接求解所导出的方程组，不仅适合定常流计算，而且适合非定常流计算。多重网格法可以有效提高计算效率，既适应显格式又适应隐格式。高阶面元法是一种较特殊的方法，涉及稠密线性方程组的求解，既可以用迭代法求解，又可以用直接法求解。

本章将分六节分别介绍显格式、多重网格法、分裂法、高阶面元法、迭代法以及稀疏带状或块带状线性代数方程组的直接求解（简称直接法）的分布式并行计算方法。多重网格法部分以隐格式计算为例。高阶面元法部分重点介绍与之相关的稠密线性方程组的并行直接求解。

8.1　显格式

任何形式的显格式，包括在空气动力学计算中广泛采用的 MacCormack 显格式，显式 TVD 格式、显式 NND 格式以及龙格—库塔方法，都可以写成如下形式：

$$Q^{n+1} = LQ^{n} \tag{8.1}$$

其中 L 为差分算子。

当流场中有边界层或激波需要分辨时，在这些地方的计算网格要局部加密，受稳定性条件的限制时间步长只能取得很小，要计算几千甚至上万步方能收敛到稳态。显格式的并行计算一般采用区域分裂（三维问题可叫空间分裂）的方法来保证各处理机之间的静态负载平衡，区域与区域之间的信息交换只发生在相邻区域之间，几乎不存在远程通信，编程简单，可扩展性好，对于处理机数目较多的 MPP 系统，并行效率也会比较高。显格式的并行计算是隐格式并行计算的基础，因为任何形式的隐格式都有系数矩阵，右端项 RHS 等需要计算，某些情况下还有人工粘性系数，物理粘性系数需要计算，而这些都是有了 n 时刻的值就可计算出来的，不需要 $n+1$ 时刻的值，因而除方程组求解以外的部分并行效率

都不会太低。

莫则尧采用有限体积离散、显式龙格—库塔迭代、两层多重网格算法在8台通过以太网连接的 SGI 工作站机群环境下对 RAE 翼身组合体跨音速绕流进行了基于 PVM 的并行计算，加速比达6.0。每台 SGI 工作站的主频为100MHz，主存32MB，数据传输速率10Mbps。

8.2 多重网格法

8.2.1 引言

当流场中存在波与波的强干扰、强烈的双曲—椭圆型混合区，且计算网格密集时，会产生极大的数值“刚性”，大大影响收敛速度，导致难以承受的计算耗时。数值实践使人们认识到，大部分用来求解离散化偏微分方程的数值迭代方法都能有效地消除高频误差分量，而在消除低频误差分量方面却差强人意。计算中正是这些顽固的低频误差严重影响了收敛速度。多重网格方法提供了一种消除低频误差的机制：当代表细网格上误差频谱的信息被传到粗网格时，细网格的低频误差分量被当作高频误差求解，即在粗网格上产生一个对细网格低频误差分量的“粗网格修正”；当上述修正被插值到细网格时，因波速一定，波长变短，从而产生容易被细网格消除的高频误差。这一从细到粗，再从粗到细的循环被称为多重网格循环。理论上这样的循环可以消除所有的误差分量，极大地加快收敛。

考虑到多重网格方法所需计算机资源的额外开销，研究其多机并行计算十分必要。如果简单地将不同的网格层计算分配给不同的机器，无法做到负载平衡，并行效率很低，实用性不大。我们研究了多重网格法与拼接网格法的匹配，实现了 PVM 平台下多机多进程多重网格计算。底部流计算所获得的高效率和可靠的结果显示了其价值。

8.2.2 多重网格原理

研究多重网格法的最初动力源于迭代求解离散的椭圆型方程时，存在各 Fourier 误差分量的不同衰减速度和衰减程度。简言之，多重网格法的逻辑是：①误差可划分为高频摆动分量和低频光滑分量；②设计某些特殊的迭代方法去消除那些摆动分量；③用多重网格来消除那些顽固的光滑分量。因此，细网格松弛迭代和粗网格修正是多重网格法的两大基础。下面结合简单的椭圆型问题，从数学上阐述其原理。

考虑一维椭圆方程。设 h 为步长，$N+1$ 为网格数。求解第一类边值问题，端点处 $n=0$ 和 $n=N$ 的值已知。方程离散后可写为：

$$-u_{j-1}+2u_j-u_{j+1}=h^2f_j=F_j \tag{8.2}$$

(8.2)的矩阵形式为：

$$Au=F \tag{8.3}$$

A 为大型稀疏矩阵。将 A 分解为 $A=D-L-U$，则雅可比松弛迭代可写为：

$$\overline{u_j^{n+1}}=D^{-1}(L+U)u_j^n+D^{-1}F_j$$

$$u_j^{n+1}=Mu_j^n+\omega D^{-1}F$$

$$M = [(1-\omega)I + \omega D^{-1}(L+U)] \tag{8.4}$$

$\|M\|$ 即为迭代法(8.4)的收敛数。迭代法收敛的充要条件是 M 的谱半径小于 1。可以验证，矩阵 M 的 $N-1$ 个特征值为：

$$\lambda_k = 1 - 2\omega\sin^2\frac{k\pi}{2N}, k = 1,2,\cdots,N-1 \tag{8.5}$$

与之对应的 N－1个独立的特征矢量为：

$$\nu_k = (\sin kx_1\pi, \sin kx_2\pi, \cdots, \sin kx_{N-1}\pi)^T, k = 1,2,\cdots,N-1 \tag{8.6}$$

其中 $x_1, x_2, \cdots, x_{N-1}$ 为网格点的坐标值。k 代表这些特征矢量的频率。k 越大，频率越高。这 $N-1$ 个特征矢量构成 $N-1$ 维线性空间，在此线性空间内的任何初始误差均可表示为：

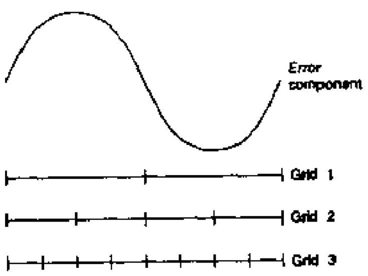
$$E^{(0)} = a_1\nu_1 + a_2\nu_2 + \cdots + a_{N-1}\nu_{N-1} \tag{8.7}$$

在迭代过程中，

$$\begin{aligned} E^{(n+1)} &= ME^{(n)} = M(ME^{(n-1)}) = \cdots \\ &= a_1\lambda_1^{n+1}\nu_1 + a_2\lambda_2^{n+1}\nu_2 + \cdots + a_{n-1}\lambda_{n-1}^{n+1}\nu_{n-1} \end{aligned} \tag{8.8}$$

显然，特征矢量代表误差分量，而对应的特征值代表误差分量的放大系数。由(8.5)知，放大系数依赖于误差频率 k 和松弛因子 ω(见图 8.1)。只有当 $0 < \omega \leqslant 1$ 时，才能保证 $|\lambda_k(M)| < 1$，雅可比松弛迭代收敛。放大系数越大，收敛越慢，反之，收敛越快。

由图 8.1 可知，对于低频分量，无论如何选取松弛因子，放大系数都不可能很小。而对于高频分量，只要妥善选择 ω，可得到较小的放大系数。因此，迭代法能将高频分量迅速衰减掉，而对那些低频分量几乎无能为力。但是，对低频和高频的区分是相对的，这与网格尺度有关。在细网格上视为低频的分量，如放到粗网格上可能被视为高频分量。图8.2给出了不同网格尺度上误差频率的区别。正是这种区别启发人们采用几重网格来消除误差分量，加快收敛。



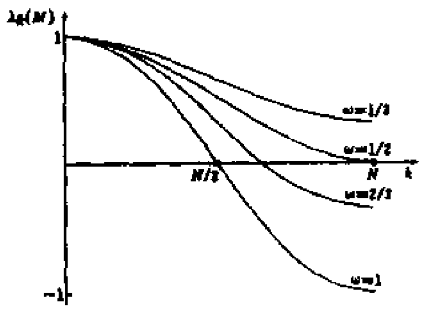


图8.1 误差分量的放大系数　　　　图8.2 不同网格尺度上误差频率的差别

多重网格计算用于求解流体力学问题，其基本思想仍然是一致的：将在不同网格尺度上的计算结合起来，利用一种网格尺度上的结果去消除另一种网格尺度上的误差分量。每一层粗网格由上一层网格间距加大一倍得到，细网格上的变量值按下式传给粗网格：

$$Q_{2h} = \frac{\sum S_h Q_h}{S_{2h}} \tag{8.9}$$

式中下标表示网格间距，Q 表示守恒变量，S 表示网格单元面积，求和在粗网格所包含的

细网格上进行。(8.9)式保证了质量、动量和能量守恒。定义了粗网格和变量值，粗网格求解过程为：

① 求解细网格，更新其变量值；

② 根据(8.9)式求粗网格的变量值；

③ 对细网格上的残差求和，得到粗网格上的残差：

$$R_{2h} = \sum R_h \tag{8.10}$$

④ 根据 R_{2h} 得到粗网格上变量的修正，并更新粗网格上的值。

下一层粗网格的求解，其网格和变量定义同前，残差由下式给出：

$$R_{4h} = \sum\left(R_{2h}^{u} + \sum R_h - R_{2h}^{t}\right) \tag{8.11}$$

其中上标 u 和 t 分别表示用更新值和转换值计算的残差。更下一层粗网格的计算依此类推。最后，用线性插值将每一层粗网格上的守恒变量增量传给上一层网格。这样，在多重网格法计算时，每一循环步最上层网格上的值得到两次修正，且每次当粗网格上的增量被插值到细网格时，产生的高频误差容易被消除，加快了收敛。

8.2.3 多重网格法与拼接网格法匹配的必要性及匹配策略

数值模拟底部流场时，底部区域收敛缓慢是较大的难题。多重网格法能有效地加快收敛，但它需要存储各网格层上的网格和物理量，所需计算机资源的额外开销大，三维计算时尤其如此，因此，对于规模较大的问题，单台微机已不能满足要求，多重网格法的多机并行处理已十分必要。如果简单地将不同的网格层计算分配给不同的机器，有如下缺点：①因不同网格层的网格数相差大于或等于1的整数倍，无法做到负载平衡，并行效率将严重下降；②随处理机数增加，只能增加网格层数，而网格层数不可能无限制地增加下去，因此，无法有效地利用较大规模的并行环境；③处理复杂底部流动困难很大，而一旦不能处理复杂外形和流场，多重网格法在计算流体力学中的实用性将大打折扣。本书研究的多重网格法与拼接网格法的匹配策略，基本克服了上述缺点，实现了 PVM 平台下多机多进程的多重网格计算。在该方法中，区域间的边界信息交换只在各区的最上一层网格之间进行，而各区的多重网格计算完全独立。这样，从用户的角度来看，多重网格与拼接网格匹配计算和拼接网格计算几乎是一致的，这种一致性使得多重网格与拼接网格相匹配策略继承了拼接网格并行算法的所有优点(如实现方便，移植性好，扩充性好等)，并保留了多重网格算法的加速机制。

下面以二维为例，讲述超声速钝头体及底部流计算时两层多重网格法与拼接网格法相匹配的策略。

计算弹身区1和底部区2两个区(见图8.3)。1区一层网格数为 $IA \times JA$，二层网格数为 $IB \times JB$；2 区一层网格数为 $IC \times JC$，二层网格数为 $ID \times JD$。计算步骤为：

① 实施1区的多重网格计算，并将一层网格上 $I1 = (IA - 1)$ 面上的数值通量插值给 2 区一层网格的 $I3 = 2$ 面，以保证通量守恒；

② 底部表面按固壁处理，修正2区一层网格中 $I3 = 2$ 到 $IC - 1$，$J3 = 1$ 到 $JC - 1$ 上的值；

③ 按(8.9)和(8.10)式计算2区二层网格中 $I4=2$ 到 $ID-1$,$J4=1$,$JD-1$ 上的守恒变量和残差,$I4=1$ 上的守恒变量和残差由1区一层网格单元 $I1=IA-3$ 和 $I1=IA-2$ 上的值提供;

④ 实施2区的多重网格计算,并将2区一层网格单元 $I3=2$ 上的守恒变量插值给1区一层网格的 $I1=IA-1$ 单元,并转①。边界处的 由新值和旧值之差获得。三维处理方法类似。

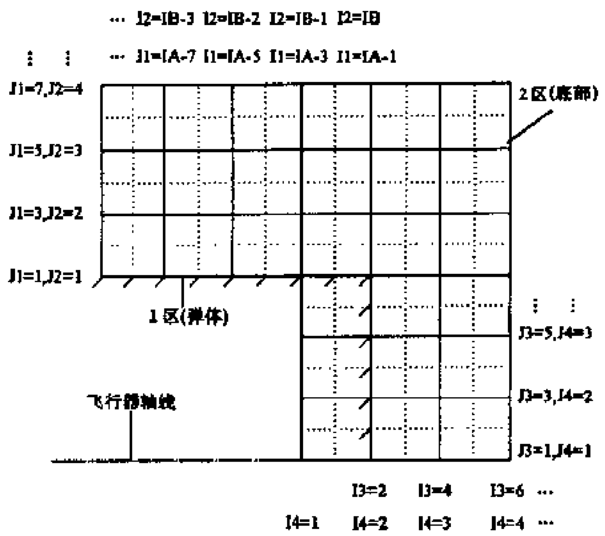


图8.3 多重网格法与拼接网络匹配示意图

根据上述多重网格法与拼接网格法匹配策略,计算了绕超声速钝头体的轴对称流场和底部流场。两区的初始网格为 $IA\times JA=82\times40$ 和 $ID\times JD=80\times86$,$M_\infty=3.5$,$Re_\infty=2.5\times10^6$。结果见表8.1(网层网格计算一步所需CPU时间大约是一层网格计算的1.2倍)。

表8.1 二维计算中多重网格法与拼接网格匹配的并行效率(3处理器)

区域 \ 效率	多重网格效率(计算步数)	多重网格效率(墙上时间)	并行效率(墙上时间)	多重网格与拼接网格匹配并行效率(墙上时间)
弹体区	470:2100	1:3.72	1:2.94	1:10.94
底部区	1200:3800	1:2.64	1:2.94	1:7.76

由表8.1可见,PVM平台下多重网格法与拼接网格法匹配算法的加速效果非常明显。

8.2.4 三维底部流场的多重网格计算

将上述多重网格法与拼接网格法相结合的方法推广到三维,实现了PVM平台下钝头体三维带尾支杆超声速底部流场的多重网格计算,弹体总网格数为80×42×41,底部总网格数为120×100×41,来流条件为 $M_\infty=3.0$,$\alpha=10^\circ$。底部由三块网格拼接在一起,以达到负载平衡。结果见表8.2。

表8.2 三维计算中多重网格法与拼接网格匹配的并行效率(4处理器)

区域 \ 效率	多重网格效率(计算步数)	多重网格效率(墙上时间)	并行效率(墙上时间)	多重网格与拼接网格匹配并行效率(墙上时间)
弹体区	440:2260	1:4.17	1:3.84	1:16.01
底部区	950:2380	1:2.03	1:3.84	1:7.79

8.2.5 结论

多重网格法与拼接网格法的匹配，使算法继承了多重网格法的加速收敛机制和拼接网格便于处理复杂流场、易于并行处理的共同优点，加速效果可观。对拼接网格并行算法的可移植性和可扩性的继承，使其在 MPP 等环境下的大规模并行计算成为可能。这对大规模数值模拟底部流场，考察各种因素对底部气动特性的影响程度，以便对实验结果进行可靠的修正，有实用价值。

8.3 分裂法

8.3.1 交替方向隐格式(ADI)和近似因子(AF)隐格式

ADI 和 AF 隐格式都可以表述成如下形式：

$$L_x L_y L_z Q^{n+1} = RHS \tag{8.12}$$

其中 L_x, L_y, L_z 是块三对角形式的局部一维算子，因而计算可以分三步进行：

$$\begin{aligned} Q^{*} &= L_x^{-1} RHS \\ Q^{**} &= L_y^{-1} Q^{*} \\ Q^{n+1} &= L_z^{-1} Q^{**} \end{aligned} \tag{8.13}$$

计算 Q^* 时，由于算子 L_x 的求逆是局部一维的，因而可以根据处理机的台数将计算空间划分成多个平行于 X 轴的长方形块(体)进行并行计算。同理，计算 Q^{**} 和 Q^{n+1} 时可以将计算空间分成多个平行于 Y 轴和 Z 轴的长方形块(体)进行并行计算。块与块或体与体之间的信息交换是局部的，远程通信很少，因而算法扩展性好，并行效率高。

日本国家航天实验室计算科学室的 T. Iwamiya 等人在数值风洞(NWT)上运行了用 AF 方法求解三维流动的 Navier-Stokes 计算程序 NS3D。NWT 由2个控制处理机(无向量部件)和140个类似 VP－400的向量处理机组成。每个向量机的时钟周期9.5纳秒(或称毫微秒)，其峰值性能为1.7GFLOPS。系统用 Cross-bar 互连网络连接而成，数据传输速率为421Mbps。采用 VP－2600超级计算机作为前端机，内存为8GB，读数据的速率为1250Mbps，写数据的速率为842Mbps。整个系统的主存为35GB，峰值运行速度为236FLOPS。

最快的速度可以通过扩大问题的规模而得到，当网格数为514×420×280时，总网格点数约为6千万，测得该系统的性能为116GFLOPS，比 NS3D 在 VP－400上所获得的最快速率快190倍。当网格为64×128×128时，问题在单个处理单元上运行不会超内存，过去在单个处理单元上需要做20小时的问题，在由64个处理单元组成的并行机上22分钟即可做完，具体结果如下：

P	2	4	8	16	32	64	128
Sp	1.89	3.72	7.49	14.96	29.77	57.65	101.7
E	0.95	0.93	0.94	0.94	0.93	0.90	0.79

8.3.2 隐式 LU 分解算法

对计算空间守恒型 Navier-Stokes 方程的无粘部分采用带加权因子 β 的显、隐混合格式，粘性部分采用显格式进行离散，得到：

$$
\begin{aligned}
&[I + \beta\Delta t(\delta_\xi A + \delta_\eta B + \delta_\zeta C)]\delta\overline{U}^n_{i,j,k} \\
&= -\Delta t[\overline{\delta}_\xi(\overline{E} - \overline{E}_v) + \delta_\eta(\overline{F} - \overline{F}_v) + \delta_\zeta(\overline{G} - \overline{G}_v)]^n_{i,j,k}
\end{aligned} \tag{8.14}
$$

其中：

$$
\delta\overline{U}^n_{i,j,k} = \overline{U}^{n+1}_{i,j,k} - \overline{U}^n_{i,j,k} \tag{8.15}
$$

Δt 为时间步长，$0 \leqslant \beta \leqslant 1$ 为控制隐式程度的参数，$\beta = 0$ 为古典显格式，$\beta = 1$ 为古典 隐格式，$\beta = 1/2$ 为 Crank-Nicholson 格式，此时格式的空间精度为二阶，其余情况格式只有一阶空间精度。等号右侧的 δ 和 $\overline{\delta}$ 分别为以网格中心和网格围线为中心的中心差分算子。为了消除可能出现的奇偶振荡和激波振荡，需要加入二阶与四阶混合人工粘性 $\overline{T}$。直接求解(8.15)需要进行七对角块矩阵的求逆，非常耗时，进行 LU 分裂得：

$$
\begin{aligned}
&[I + \beta\Delta t(\delta^-_\xi A_1 + \delta^-_\eta B_1 + \delta^-_\zeta C_1)][I + \beta\Delta t(\delta^+_\xi A_2 + \delta^+_\eta B_2 + \delta^+_\zeta C_2)]\delta\overline{U}^n_{i,j,k} \\
&= -\Delta t[\overline{\delta}_\xi(\overline{E} - \overline{E}_v) + \delta_\eta(\overline{F} - \overline{F}_v) + \delta_\zeta(\overline{G} - \overline{G}_v) + \overline{T}]^n_{i,j,k}
\end{aligned} \tag{8.16}
$$

其中：

$$
\begin{aligned}
A_1 &= \frac{(A + r_A I)}{2}, A_2 = \frac{(A - r_A I)}{2} \\
B_1 &= \frac{(B + r_B I)}{2}, B_2 = \frac{(B - r_B I)}{2} \\
C_1 &= \frac{(C + r_C I)}{2}, C_2 = \frac{(C - r_C I)}{2}
\end{aligned} \tag{8.17}
$$

r_A、r_B、r_C 分别为矩阵 A、B、C 的谱半径，δ^+ 和 δ^- 为一阶向前、向后差分算子，I 是单位矩阵。显然 A_1、B_1、C_1 的特征值非负，A_2、B_2、C_2 的特征值非正。

向下扫描：

$$
\begin{aligned}
&[I + \beta\Delta t(A_1 + B_1 + C_1)]\delta\overline{U}^*_{i,j,k} \\
&= -\Delta t[\overline{\delta}_\xi(\overline{E} - \overline{E}_v) + \delta_\eta(\overline{F} - \overline{F}_v) + \delta_\zeta(\overline{G} - \overline{G}_v) + \overline{T}]^n_{i,j,k} \\
&\quad + \beta\Delta t(A_1\delta\overline{U}^*_{i-1,j,k} + B_1\delta\overline{U}^*_{i,j-1,k} + C_1\delta\overline{U}^*_{i,j,k-1})
\end{aligned} \tag{8.18}
$$

向上扫描：

$$
[I - \beta\Delta t(A_2 + B_2 + C_2)]\delta\overline{U}^*_{i,j,k} + \beta\Delta t(A_1\delta\overline{U}^*_{i-1,j,k} + B_1\delta\overline{U}^*_{i,j-1,k} + C_1\delta\overline{U}^*_{i,j,k-1}) \tag{8.19}
$$

$$
= \delta\overline{U}^*_{i,j,k} - \beta\Delta t(A_2\delta\overline{U}^n_{i+1,j,k} + B_2\delta\overline{U}^n_{i,j+1,k} + C_2\delta\overline{U}^n_{i,j,k+1})
$$

经相似变换可将 $A + B + C$ 对角比，即存在相似变换矩阵 P 使得

$$P^{-1}(A+B+C)P=\Lambda \tag{8.20}$$

其中 Λ 为对角阵，对角线上的元素为

$$\begin{aligned}
&\lambda_{11}=\lambda_{22}=\lambda_{33}=U+V+W\\
&\lambda_{44}=U+V+W-c\sqrt{l_1^2+l_2^2+l_3^2}\\
&\lambda_{55}=U+V+W-c\sqrt{l_1^2+l_2^2+l_3^2}
\end{aligned} \tag{8.21}$$

$$\begin{pmatrix}U\\V\\W\end{pmatrix}=\begin{pmatrix}\xi_x & \xi_y & \xi_z\\ \eta_x & \eta_y & \eta_z\\ \zeta_x & \zeta_y & \zeta_z\end{pmatrix}\begin{pmatrix}u\\v\\w\end{pmatrix} \tag{8.22}$$

$$\begin{aligned}
l_1&=\xi_x+\eta_x+\zeta_x\\
l_2&=\xi_y+\eta_y+\zeta_y\\
l_3&=\xi_z+\eta_z+\zeta_z
\end{aligned} \tag{8.23}$$

c 是当地音速，当地时间步长由下式估计得到

$$\Delta t=\frac{C_n}{(r_A+r_B+r_C)} \tag{8.24}$$

其中 C_n 为 Courant 数，利用(8.18)、(8.19)和(8.20)可以转化为解耦的单个方程 ($m=1$，$2,\cdots,5$)。

往下扫描：

$$\begin{aligned}
&(P^{-1}\delta\overline{U}^{*}_{i,j,k})_m\\
&=[-\Delta tP^{-1}((\overline{\delta}_\xi(\overline{E}-\overline{E}_v)+\delta_\eta(\overline{F}-\overline{F}_v)+\delta_\zeta(\overline{G}-\overline{G}_v)+\overline{T})^n_{i,j,k}\\
&\quad-\beta(A_1\delta\overline{U}^{*}_{i-1,j,k}+B_1\delta\overline{U}^{*}_{i,j-1,k}+C_1\delta\overline{U}^{*}_{i,j,k-1}))]_m/((1+\frac{\beta C_n}{2})I+\frac{\beta\Delta t}{2}\Lambda)_m
\end{aligned} \tag{8.25}$$

往上扫描：

$$\begin{aligned}
&(P^{-1}\delta\overline{U}^{n}_{i,j,k})_m\\
&=[P^{-1}(\delta\overline{U}^{*}_{i,j,k}-\beta\Delta t(A_2\delta\overline{U}^{n}_{i+1,j,k}+B_2\delta\overline{U}^{n}_{i,j+1,k}+C_2\delta\overline{U}^{n}_{i,j,k+1}))]_m/((1+\frac{\beta C_n}{2})I-\frac{\beta\Delta t}{2}\Lambda)_m
\end{aligned} \tag{8.26}$$

且

$$\overline{U}^{n+1}_{i,j,k}=P(P^{-1}\delta\overline{U}^{n}_{i,j,k})+Q^{n}_{i,j,k} \tag{8.27}$$

这样七对角块矩阵的求逆变成了一个标量除法，非常接近显格式的求解，所不同的是右端项中有扫描过的刚计算出的变化量马上需要使用，当个别的变化量不在本区时需要用到相邻区 n 时刻的变化量的值而不是相邻区刚算出来的值，因而残差有些细微的差别，对收敛速度的影响需要引起重视。

我们曾用 LU 方法对 Navier-Stokes 方程进行计算，求稳态解，数值实验是在由 2 台 486 微机组成的分布式并行模拟环境下进行，计算发现并行计算的残差与串行计算的残差略有差别，单机串行计算 654 步收敛，总时间为 T=21 058.8 秒，两机二进程并行计算 650 步就达到了收敛，运行时间为 T=11 050.0 秒，加速比为 1.94。之后，我们在国产 MPP 机上用 LU 方法并行求解了多类复杂流场的流动问题，均能得到收敛解，且加速比

很高。

8.4　高阶面元法

高阶面元法的出发方程为：

$$(1 - M_\infty^2)\varphi_{xx} + \varphi_{yy} + \varphi_{zz} = 0 \tag{8.28}$$

令

$$s = \mathrm{sign}(1 - M_\infty^2), \beta = \sqrt{s(1 - M_\infty^2)} \tag{8.29}$$

将坐标系作适当的缩放：

$$\begin{aligned} \bar{x} &= x \\ \bar{y} &= \beta y \\ \bar{z} &= \beta z \end{aligned} \tag{8.30}$$

则方程(8.28)可以写成：

$$s\varphi_{\bar{x}\bar{x}} + \varphi_{\bar{y}\bar{y}} + \varphi_{\bar{z}\bar{z}} = 0 \tag{8.31}$$

当 $M_\infty = 0$ 或 $M_\infty = \sqrt{2}$ 时，(8.31)与(8.28)相同。当 $M_\infty < 1$ 时，流动为亚声速流动，此时 $s = 1$，(8.31)为拉普拉斯方程；当 $M_\infty > 1$ 时，流动为超声速流动，此时，$s = -1$，(8.31)为双曲型方程。应用格林定理，可以将(8.31)转化成沿物面的积分方程：

$$\varphi(P) = -\frac{1}{K}\iint_S \left[\frac{\sigma}{R} - \mu \boldsymbol{n} \cdot \nabla\left(\frac{1}{R}\right)\right] \mathrm{d}S \tag{8.32}$$

其中：σ：源密度

μ：偶极子密度

$\boldsymbol{n}$：物面的单位外法向矢量

R：P 点至在曲面 S 上的积分点的距离

K：常数，超音速时取 2π，亚音速时取 4π

将复杂飞行器表面分成若干个网络，每个网络由 $M \cdot N$ 个网格点组成，形成 $(M-1) \cdot (N-1)$ 个面元，将网格中心及各连中点互连构成 8 个三角形子面元，然后布置线性源和二次偶极子，考虑边界条件和偶极子匹配条件，(8.32)式的计算可以转化为一个高阶稠密线性方程组的求解，即下式 [AIC] 的求逆问题：

$$[AIC][\lambda] = [b] \tag{8.33}$$

当攻角及边界条件变化时，[b] 改变但[AIC]不变，因面一次求解可以带多个右端项，求解多组状态。当矩阵[AIC]的阶数达到上千时，在单机上求解，一是容易超内存，二是求逆的时间太长，满足不了工程实践的需要，因此必须寻求该问题的并行算法。基于列交换的 Gauss-Jordan 并行求逆算法能满足此需要。该方法采用块循环的数值分配方式，有效地减少了远程通信，负载平衡性较好，对超立方体结构并行机尤其适合。

在高阶稠密矩阵的并行求逆过程中，由于数据相关性强，如何分配数据将影响负载平衡性及通信开销，对算法的性能及可扩展性有较大的影响。算法是面向体系结构的，因此数据分配方案与具体应用的体系结构密切相关，这里给出的算法是面向超立方体结构的

并行机系统，采用的是块循环数据分配方式，即将 $n\times n$ 矩阵分解成大小为 $nr_1\times nr_2$ 的块，行列方向固定距离的元素被分配给同一个处理机，行列方向的步距分别为 r_1 和 r_2，其中，$r_1\times r_2=p$，组成了 $r_1\times r_2$ 处理机网格。nr_1、nr_2 与 r_1、r_2 的关系分别为：

$$nr_1=n/r_1,nr_2=n/r_2 \tag{8.34}$$

将矩阵 A 中行号为 $(k-1)/r_2+r_1\times(i-1)+1$，列号为 $\mathrm{mod}(k-1,r_2)+r_2\times(j-1)+1$ 的元素分配给处理机 p_k，作为 p_k 的元素 $mma(i,j)(i=1,2,\cdots,nr_1,j=1,2,\cdots,nr_2,k=1,2,\cdots,p)$。例如，当 $n=6,p=6,r_1=2,r_2=3$ 时，数据分配方式如下：

$$\begin{bmatrix} p_1 & p_2 & p_3 & p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 & p_4 & p_5 & p_6 \\ p_1 & p_2 & p_3 & p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 & p_4 & p_5 & p_6 \\ p_1 & p_2 & p_3 & p_1 & p_2 & p_3 \\ p_4 & p_5 & p_6 & p_4 & p_5 & p_6 \end{bmatrix}$$

设 rp_1 和 rp_2 分别是所有不小于 r_1 和 r_2 且是 2 的幂的最小值，$d_1=\log_2(rp_1)$，$d_2=\log_2(rp_2)$，则处理机 $p_{me}(me=1,2,\cdots,p)$（处理机网格中与处理机 p_{me} 处于同一行或同一列的处理机分别记录在 $rowp$ 和 $colp$ 中）的算法描述如下：

for k =1 to n do

　　if p_{me} 上有第 K 行元素 then

　　　这些元素构成局部主元行 $pr(i),i=1,2\cdots,nr_1$；将 pr 广播给 $p\in rowp$.

　　else 接收消息到 $pr(i),i=1,2\cdots,nr_1$

　　end if

　　取 pr 中绝对值最大的元素为 pe，相应列号为 col，且记录 p_{me} 上有关第 col 列元素段到 $pc(i),i=1,\cdots,nr_2$.

　　if p_{me} 上有第 K 行元素 and $col\neq K$ then

　　　　交换这两列元素

　　end if　（可能有交换）

　　在 colp 中的 nr_2 个处理机上比较 pe. 最终，每个处理机得到全局主元 pe 和相应的 pc. 用 pc、pr、pe 修改本地数据 $mma(i,j),i=1,2,\cdots,nr_1,j=1,2,\cdots,nr_2$

end for

设一个整型数的消息长度为 INT，一个浮点数的消息长度为 $FLOAT$。该算法的第 k 次循环要做三项工作：

(1) 在处理机网格的列中广播具有 A 的第 k 行元素的主元消息 $pr(j),(j=1,2,\cdots,nr_2)$，然后每个处理机自选决定局部主元 $pe1$、该主元所在的列 $pc(i),(i=1,2,\cdots,nr_1)$ 以及相应的列号为 col。这里 r_2 个处理机具有主元行消息，分别向 r_1-1 个处理机广播该消息，消息长度为 $nr_2\times FLOAT$，这 r_2 个处理机是并行广播的。

(2) 在处理机两格的同一行中 r_2 个处理机比较各自的 pel，得到最大值即全局主元 pe，相应的主元列 $pc(i),(i=1,2,\cdots,nr_1)$ 以及列号 col。每个处理机 p_{me} 依次与 p_j 相互比较，其中 $j-1$ 与 $me-1$ 的二进制中某一位互为相反，这样的 j 至多有 d_2 个，即每个处理

机至多要与 d_2 个处理机进行点对点通信，消息长度为 $1 \times INT + (1 + nr_1) \times FIOAT$。

(3) 用 pe，pr 和 pc 修改本地数据 $mma(i,j)$，$i = 1,2,\cdots,nr_1$，$j = 1,2,\cdots,nr_2$，该部分不与其他处理机通信，为计算核心块，计算量为 $O(nr_1 * nr_2)$。

所用算法是针对超立方结构的并行机提出来的，这是因为它的通信均是在处理机 i 和 j 间进行，其中 i 和 j 的二进制数的某一位是互反的，在超立方体结构的计算机上，具有这种性质的处理机存在直接的物理连接，通信所需的时间少。文中 r_1、r_2 有一个隐含限制，即它们是 2 的幂，则 p 也是 2 的幂，本文通过引进 rp_1、rp_2 可消除此限制，同时在步(1)中利用广播方式，而不是采用的点对点通信，减少了等待时间。

高阶稠密矩阵求逆的数值实验在由 4 台 SGI 工作站组成的多机系统上完成，每台工作站的性能参数网络传输率与显格式求解时一样。其计算结果如下：

N	$T_1(s)$	$T_p(s)$	S_p	E
500	81.9	52.9	1.55	38.7%
800	281.9	123.0	2.28	57.0%
1000	532.7	204.5	2.56	64.0%

表中当 $n = 500$ 时，S_p 和 E 不理想，原因在于此时问题的规模还不够大，通信量与计算量相比所占比重较大。当 $n = 1000$ 时 4 处理机加速比达 2.56，仍然不高，这是因为这种算法是针对具有超立方体互连网络的并行机提出的。

8.5 迭代法

对于中等规模的线性方程组来说，直接法求解工作量少，精度高；但对大型稀疏线性方程组而言，直接法不能充分利用方程组系数矩阵的稀疏特性来有效减少存储量，降低计算开销，而且并行难度大。而迭代求解法在这方面却有优势，它将问题转化为构造一个无穷迭代序列来逐步逼近方程组的解。跟直接法比较，迭代法具有方法简单、所需空间小、易并行等优点，因而在大规模计算中呈现出优势。

现有的针对一般形状线性方程组的迭代法很多，如古典的 Jacobi 迭代、Gausss-Seidel 迭代、SOR 迭代、SSOR 迭代等，它们稍加变形就可以适用于稀疏带状线性方程组，但这些迭代法一般收敛很慢。迭代法中还有 Krylov 子空间迭代法，它具有计算量和存储量少、易于并行等优点。Krylov 迭代法一般要结合预条件子，找到好的预条件子是关键问题，但对有些问题离散出来的线性系统，难于找到好的预条件于，或者计算好的预条件子的时间很长。总的说来，对迭代法，国内外已经有了较多的研究。我们这里主要介绍古典迭代法。

8.5.1 迭代法的基本理论

一、迭代法的一般形式

用迭代法解 n 阶线性方程组 $Ax = b$ 的一般形式是

$$x^{(k)} = F_k(A,b,x^{(0)},x^{(1)},\cdots,x^{(k-1)}) \qquad k = 1,2,\cdots \tag{8.35}$$

对任意给定的初始向量 $x^{(0)}$，要求 $\lim_{k\to\infty} x^{(k)} = x$。选择不同的 F_k，即可以得到各种不同的迭代公式。如果 F_k 只依赖于 $x^{(k-1)}, x^{(k-2)}, \cdots, x^{(k-p)}$，则称 F_k 为 p 阶迭代公式。实践中，p 取的比较小。当 $p = 1$ 时，称为一阶迭代公式

$$x^{(k)} = F_k(A, b, x^{(k-1)}) \qquad k = 1, 2, \cdots \tag{8.36}$$

当 F_k 与迭代次数无关时，称为定常迭代。最简单的一阶迭代是 $x^{(k-1)}$ 的线性函数

$$x^{(k)} = H_k x^{(k-1)} + g_k \qquad k = 1, 2, \cdots \tag{8.37}$$

式中 $H_k = H_k(A, b)$ 是 $n \times n$ 方阵，g_k 是含 n 个分量的列向量。在定常情况下，$H_k = H$，$g_k = g$。于是

$$x^{(k)} = Hx^{(k-1)} + g \qquad k = 1, 2, \cdots \tag{8.38}$$

可以证明，迭代法的本质是将系数矩阵分解为两个矩阵的和，不同的分解方式和原则构成不同的迭代方法，这跟直接法形成鲜明对照，后者是将系数矩阵分解为矩阵的乘积。

二、迭代法的收敛性

将 $Ax = b$ 的解 x_T 代入一阶迭代格式

$$x^{(k)} = H_k x^{(k-1)} + g_k \tag{8.39}$$

得到

$$x_T = H_k x_T + g_k \tag{8.40}$$

由(8.39)～(8.40) 并令 $e^{(k)} = x^{(k)} - x_T$，得

$$e^{(k)} = H_k e^{(k-1)} \tag{8.41}$$

(8.41)式说明，误差向量 $e^{(k)}$ 满足(8.39)式的齐次式，这是线性迭代的基本特征。于是

$$e^{(k)} = H_k H_{k-1} \cdots H_1 e^{(0)} \tag{8.42}$$

令 $M_k = H_k H_{k-1} \cdots H_1$，对任意初值 $x^{(0)}$，误差 $\lim_{k\to\infty} e^{(k)} = 0$ 的充分必要条件是 $\lim_{k\to\infty} M_k = 0$。在定常迭代(包括循环迭代) 的情况下，对任意初始向量 $x^{(0)}$，迭代收敛的充分必要条件是矩阵 H 的谱半径 $\rho(H) < 1$。

三、迭代法的收敛速度

对于收敛的定常迭代，有下列形式

$$x^{(k)} - x_T = H(x^{(k-1)} - x_T) \tag{8.43}$$

根据矩阵、向量模的性质知，存在常数 c，$0 \leqslant c < 1$，使得

$$\| x^{(k)} - x_T \|_\infty \leqslant c \| x^{(k-1)} - x_T \|_\infty \tag{8.44}$$

通常 $c \approx \rho(H)$。实践中为了估计常数 c，可利用恒等式

$$x^{(k+1)} - x^{(k)} = H(x^{(k)} - x^{(k-1)}) \tag{8.45}$$

取比较大的 m，选用

$$c_k = \| x^{(k+1)} - x^{(k)} \|_\infty / \| x^{(k)} - x^{(k-1)} \|_\infty \tag{8.46}$$

$$k = m, m+1, \cdots, m+l-1$$

$$c = \sum_{k=m}^{m+l-1} c_k / l \tag{8.47}$$

多数情况下，这样估计 c 是比较准确的。如果

$$\| x^{(k+1)} - x^{(k)} \|_\infty \geqslant \| x^{(k)} - x_T \|_\infty - \| x^{(k-1)} - x_T \|_\infty$$

$$\geqslant \frac{1}{c} \| x^{(k+1)} - x_T \|_\infty - \| x^{(k+1)} - x_T \|_\infty \tag{8.48}$$

$$= \frac{1-c}{c} \| x^{(k+1)} - x_T \|_\infty$$

近似成立，则可以用下列估计式

$$\| x^{(k+1)} - x_T \|_\infty \leqslant \frac{c}{1-c} \| x^{(k+1)} - x^{(k)} \|_\infty \tag{8.49}$$

对于给定的误差限 ε，为了保证 $\| x^{(k+1)} - x_T \|_\infty < \varepsilon$，可以取

$$\| x^{(k+1)} - x^{(k)} \|_\infty \leqslant \frac{1-c}{c} \varepsilon \tag{8.50}$$

当 c 接近 1 时，收敛十分缓慢。即使 $\| x^{(k+1)} - x^{(k)} \|_\infty$ 很小，也不能保证 $\| x^{(k+1)} - x_T \|_\infty < \varepsilon$。为了使 $\| x^{(k+1)} - x_T \|_\infty < \varepsilon$，必须使用(8.50)式来控制迭代过程。

迭代法的运算量与初值 $x^{(0)}$ 相关，为了使 $\| x^{(k)} - x_T \|_\infty < \delta \| x^{(0)} - x_T \|_\infty$，利用

$$\| x^{(k+1)} - x_T \|_\infty \approx c \| x^{(k)} - x_T \|_\infty \tag{8.51}$$

有

$$\| x^{(k)} - x_T \|_\infty \approx c^k \| x^{(0)} - x_T \|_\infty \tag{8.52}$$

于是 $c^k \approx \delta$，即 $k \approx \ln\delta / \ln c$。

8.5.2 几种简单的迭代法及其并行

一、Jacobi 迭代

对线性方程组 $Ax = b$，将系数矩阵 A 分解为下列形式

$$A = D + L + U \tag{8.53}$$

其中 D 为对角线矩阵，L 和 U 分别是对角线元素为零的下三角矩阵和上三角矩阵。用 $D + L + U$ 代替矩阵 A，(8.53)式可以改写成

$$Dx = -(L+U)x + b \tag{8.54}$$

当矩阵 A 的对角元素非零时，可建立迭代公式

$$x_i^{(k)} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j^{(k-1)} + \frac{b_i}{a_{ii}} \quad i = 1,2,\cdots,n; k = 1,2,\cdots \tag{8.55}$$

写成矩阵形式为

$$x^{(k)} = D^{-1}(L+U)x^{(k-1)} + D^{-1}b \quad k = 1,2,\cdots \tag{8.56}$$

公式(8.56)称为 Jacobi 迭代公式，它的迭代矩阵

$$H_j = -D^{-1}(L+U) \tag{8.57}$$

当矩阵 A 非奇异且对角线上存在零元素时，通过矩阵的行列互换，总可以保证 D 非奇异。

当 $\rho(H_j) < 1$ 时，迭代收敛。一个实用的判定 Jacobi 迭代收敛的充分条件是矩阵 A 的对角元素占优。

由公式(8.56)可以看出，Jacobi 迭代的每一步的值只依赖于上一迭代步的值，具有很高的并行度。对稀疏线性方程组，更有交换的数据量小的特点，我们下面以稀疏带状五对

角线性方程组的并行 Jacobi 迭代法为例说明。

对方程组 $Ax=b$, A 为如下矩阵：

$$
\begin{bmatrix}
d_1 & g_1 & & & h_1 & & & & \\
e_2 & d_2 & g_2 & & & h_2 & & & \\
 & e_3 & d_3 & g_3 & & & \ddots & & \\
 & & & & & & & h_{im\times jm-jm-1} & \\
f_{jm+1} & & & \ddots & \ddots & \ddots & & & h_{im\times jm-jm} \\
 & f_{jm+2} & & & \ddots & \ddots & \ddots & & \\
 & & \ddots & & & & & & \\
 & & & f_{im\times jm-1} & & & e_{im\times jm-1} & d_{im\times jm-1} & g_{im\times jm-1} \\
 & & & & f_{im\times jm} & & & e_{im\times jm} & d_{im\times jm}
\end{bmatrix}
$$

迭代公式为

$$
x_i^{(k)} = \frac{b_i}{d_i} - \frac{f_i}{d_i}x_{i-jm}^{(k-1)} - \frac{e_i}{d_i}x_{i-1}^{(k-1)} - \frac{g_i}{d_i}x_{i+1}^{(k-1)} - \frac{h_i}{d_i}x_{i+jm}^{(k-1)}
$$

$$
i = jm+1, jm+2, \cdots, im\times jm-1; \quad k=1,2,\cdots \tag{8.58}
$$

对 $x_i^{(k)}$, $1\leqslant i\leqslant jm$ 或 $im\times jm - jm\leqslant i\leqslant im\times jm$，迭代计算公式稍有差异。

设使用 p 台处理机，按行平均分配存储五条对角线上的元素和右端项，需要同时存储 $x^{(k)}$ 和 $x^{(k-1)}$，各处理机上不仅要存储本机的各 $im\times jm/p$ 个 $x^{(k)}$ 和 $x^{(k-1)}$，还要存储邻机上的各 jm 个(对首尾两台处理机)或各 $2\times jm$ 个 $x^{(k)}$ 和 $x^{(k-1)}$(对中间的处理机)。求解过程可以简略地描述如下：

Step1：

各处理机以某种方法求得本机上的初值 $x^{(k-1)}$, $k=1$，并交换交界处的 $x^{(k-1)}$；

Step2：

各处理机并行地利用(8.58)式由 $x^{(k-1)}$ 求本机上的 $x^{(k)}$；

Step3：

以某种方法求出迭代误差。若迭代误差已经达到要求，则结束；否则，各处理机交换交界处的 $x^{(k)}$, k 增加 1，转 Step2。

可见，Jacobi 迭代容易并行且平均并行度高。

二、Gauss-Seidel 迭代

Gauss-Seidel 迭代的计算公式是

$$
x_i^{(k)} = -\frac{1}{a_{ii}}\left(\sum_{j=1}^{i-1} a_{ij}x_j^{(k)} + \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)} - b_i\right) \tag{8.59}
$$

$$
i=1,2,\cdots,n; \quad k=1,2,\cdots
$$

写成矩阵形式是

$$
x^{(k)} = -D^{-1}(Lx^{(k)} + Ux^{(k-1)}) + D^{-1}b \tag{8.60}
$$

于是

$$
x^{(k)} = -(D+L)^{-1}Ux^{(k-1)} + (D+L)^{-1}b \tag{8.61}
$$

因此，Gauss-Seidel 迭代相当于矩阵分解

$$A = (D + L) + U \tag{8.62}$$

迭代矩阵

$$H = -(D + L)^{-1}U \tag{8.63}$$

$$g = (D + L)^{-1}b \tag{8.64}$$

Gauss-Seidel 迭代从第 i 个方程计算 $x_i^{(k)}$ 时，立即用前 $i-1$ 个方程计算出来的 $x_1^{(k)}$、$x_2^{(k)}, \cdots, x_{i-1}^{(k)}$ 代替 $x_1^{(k-1)}, x_2^{(k-1)}, \cdots, x_{i-1}^{(k-1)}$，不仅可以节省一组存储单元，而且还能加快迭代收敛速度。

两个判定 Gauss-Seidel 迭代收敛的充分条件是矩阵 A 有对角优势或矩阵 A 正定。

对 Gauss-Seidel 迭代，直接按(8.59)式并行计算需要大量的通信与同步操作，开销极大。

我们可以通过采用对节点进行红、黑排序的方法来实现 Gauss-Seidel 迭代的并行计算。如 Poisson 方程

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f(x,y) \tag{8.65}$$

的五点差分近似

$$\frac{u_{i+1,j} + u_{i-1,j} - 2u_{i,j}}{h_1^2} + \frac{u_{i,j+1} + u_{i,j-1} - 2u_{i,j}}{h_2^2} = -f_{i,j} \tag{8.66}$$

对上述差分方程进行数值离散后，可得稀疏带状五对角线性方程组。这里我们考虑单个节点的迭代修正。

对于单位正方形网格 ($h_1 = h_2 = 1$)，上式化简为

$$u_{i,j} = \frac{1}{4}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} + f_{i,j}) \tag{8.67}$$

使用 Gauss-Seidel 迭代公式，得

$$u_{i,j}^{(k+1)} = \frac{1}{4}(u_{i,j-1}^{(k+1)} + u_{i-1,j}^{(k+1)} + u_{i,j+1}^{(k)} + u_{i+1,j}^{(k)} + f_{i,j}) \tag{8.68}$$

采用红—黑排序策略，将节点分成红、黑两个子集。当 $i+j$ 为偶数时，节点(i,j) 记为红点；当 $i+j$ 为奇数时，节点(i,j) 记为黑点，这样红点的邻接点仅为黑点。Gauss-Seidel 迭代的并行计算步骤为

Step1：

各处理机以某种方式得到初值 $u^{(k)}, k = 0$；

Step2：

各处理机并行地对 $i+j$ 为偶数的红点计算

$$u_{i,j}^{(k+1)} = \frac{1}{4}(u_{i,j-1}^{(k)} + u_{i-1,j}^{(k)} + u_{i,j+1}^{(k)} + u_{i+1,j}^{(k)} + f_{i,j}) \tag{8.69}$$

Step3：

各处理机交换边界数据后，并行地对 $i+j$ 为奇数的黑点计算

$$u_{i,j}^{(k+1)} = \frac{1}{4}(u_{i,j-1}^{(k+1)} + u_{i-1,j}^{(k+1)} + u_{i,j+1}^{(k+1)} + u_{i+1,j}^{(k+1)} + f_{i,j}) \tag{8.70}$$

注意右端的四个 u 值为红点上的值，已经由(8.69)式算出，因此这一步可以并行计算

所有黑点的修正值。

Step4:

若误差已经满足要求,则停止;否则,各处理机交换边界数据,k 增加 1,转 Step2。

可以证明使用(8.69)、(8.70)式组成的迭代过程,其收敛速度跟使用(8.68)式的相同。

三、逐次超松弛(SOR)迭代法

SOR 迭代是一种进一步提高收敛速度的方法。记 $u_G^{(k)}$ 是 Gauss-Seidel 迭代法的第 k 次迭代值,SOR 算法的迭代公式为

$$u_{i,j}^{(k+1)} = u_{i,j}^{(k)} + \omega[u_{i,j(G)}^{(k+1)} - u_{i,j}^{(k)}] \tag{8.71}$$

式中 ω 称为松弛因子。

可以证明当 $\omega < 2$ 时,(8.71)式是收敛的。$\omega > 1$ 称为超松弛,$\omega < 1$ 称为低松弛,$\omega = 1$ 即为 Gauss-Seidel 迭代法。选择不同的 ω,(8.71)式的收敛速度不同。

将(8.68)式代入(8.71)式,得

$$u_{i,j}^{(k+1)} = (1-\omega)u_{i,j}^{(k)} + \frac{1}{4}\omega(u_{i,j-1}^{(k+1)} + u_{i-1,j}^{(k+1)} + u_{i,j+1}^{(k)} + u_{i+1,j}^{(k)} + f_{i,j}) \tag{8.72}$$

使用红—黑排序,同样可以方便地对(8.72)式进行并行计算。

四、线迭代法

以上介绍的几种方法仅考虑对单个节点的迭代修正,而线迭代方法则是考虑一条网格线上全部节点的迭代修正。

线 Jacobi 迭代公式

$$u_{i-1,j}^{(k+1)} - 4u_{i,j}^{(k+1)} + u_{i+1,j}^{(k+1)} = -f_{i,j} - u_{i,j+1}^{(k)} - u_{i,j-1}^{(k)} \tag{8.73}$$

上式的左端表明,完成第 j 条网格线的一次迭代要求解一个三对角线方程组,可利用追赶法求解。若求解的 Poisson 方程差分问题的规模为 $n \times n$,则 n 个三对角线方程组可并行求解。可证明,线 Jacobi 迭代法的收敛速度是点 Jacobi 迭代法收敛速度的两倍。

若在(8.73)式中用 $u_{i,j-1}^{(k+1)}$ 替换 $u_{i,j-1}^{(k)}$,则得到线 Gauss-Seidel 迭代法计算公式。在并行机上,将网格线分成红线和黑线两组,首先利用黑线上前次迭代值计算红线上的迭代值,全部红线的值可以同时计算;然后利用红线已迭代修正的值计算黑线上的迭代值,所有黑线上的值可同时计算。

按下述步骤可以构造线超松弛迭代法的迭代公式。首先计算

$$u_{i-1,j(G)}^{(k+1)} - 4u_{i,j(G)}^{(k+1)} + u_{i+1,j(G)}^{(k+1)} = -f_{i,j} - u_{i,j+1}^{(k)} - u_{i,j-1}^{(k)} \tag{8.74}$$

然后对 $u_{i,j}^{(k)}$ 替换 $u_{i,j(G)}^{(k+1)}$ 线性插值

$$u_{i,j}^{(k+1)} = (1-\omega)u_{i,j}^{(k)} + \omega u_{i,j}^{(k+1)} \tag{8.75}$$

线超松弛迭代法的并行计算与线 Gauss-Seidel 迭代法类似。

8.6 直接法

8.6.1 三对角和块三对角方程组的直接并行计算

在 CFD 中,对一维问题进行三点中心差分离散,可以得到三对角方程组(不可压流

动，非耦合求解时）或者块三对角方程组（可压流动或不可压流动，耦合求解时）。对二维和三维问题，经过控制方程进行近似因子分裂（AF）、交替方向隐格式（ADI）或者线松弛方法进行处理后，通常可转化为三对角或者块三对角方程组的求解。所以三对角和块三对角方程组的求解是 CFD 计算中经常碰到的问题，其算法经过推广可以求解非稀疏（块）五、七对角方程组。

现在三对角方程组的直接并行求解方法比较多，如循环约化法、Stone 算法、Wang 提出的分裂法、P. H. Michielse 和 H. Van der Vorst 提出的 Michielse&Vorst 算法等。近年来迟利华提出了双向并行分裂(DPP)法，其通信建立次数为 $2(p-1)$，通信量为 $3(p-1)$，通过改变消元次序，减少了等待，比 Wang 的分裂法和 Pichielse&Vors 算法有改进。我们也在考虑计算后有结果传输和数据交换的情况下，提出了 SPP 算法。这些算法都采用分治的策略，首先由各处理机对子系统进行消元，然后求解缩减系统（中间需要通信），最后将缩减系统的解回代到各子系统求出结果。

注意我们这里讨论的并行算法主要是面向分布存储的并行机系统的，对其他类型的并行机系统，则算法的描述需要修改，但算法思想是相通的。

下面着重介绍 SPP 算法。

一、SPP 算法

1. 算法描述

三对角方程组 $Ax = b$ 表示如下：

$$\begin{bmatrix} d_1 & c_1 & & & & \\ a_2 & d_2 & c_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & d_{n-1} & c_{n-1} \\ & & & & a_n & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix}$$

设处理机数为 p，不失一般性，设 $n = p \times k$，将处理机编号为 $N_1, N_2, \cdots, N_p$。系数矩阵 A 和右端项 b 被预先按顺序平均分配到各处理机上，即各处理机分别分配系数矩阵中连续的 k 行和右端项 b 的 k 个连续的分量，编号越小的处理机分配的数据的下标越小。消元过程中引入的新的非零元就存储在 a_i 和 c_i 中，$i = 1, \cdots, n$。算法描述如下：

Step1：

各处理机读取各自的数据。

Step2：

对 N_1，首先将第一行对角元化为 1，然后自上往下用第 $i-1$ 行消去本机上的下次对角元 a_i，$2 \leqslant i \leqslant k$，消去每一行的下次对角元后将该行的对角元化为 1。对 N_j，$2 \leqslant j \leqslant p-1$，首先自上往下用第 $i-1$ 行消去第 i 行的下次对角元 a_i，$(j-1)k+2 \leqslant i \leqslant jk$，消去每一行的下次对角元之后将该行的对角元化为 1，再自下往上用第 $i+1$ 行消去第 i 行的上次对角元 c_i，$(j-1)k+1 \leqslant i \leqslant jk-2$。对 N_p，首先将最后一行的对角元化为 1，然后自下往上用第 $i+1$ 行消去第 i 行的上次对角元 d_i，$n-k+1 \leqslant i \leqslant n-1$，消去每一行的上次对角元之后将该行的对角元化为 1。经过 step2 后，各处理机上的系数矩阵形状如下（图

中略去了一些元素全为零的块,空白处为零元素,下同):

$$
\begin{bmatrix} 1 & * & & & \\ & 1 & * & & \\ & & \ddots & \ddots & \\ & & & 1 & * \\ & & & & 1 & * \end{bmatrix}
\quad
\begin{bmatrix} * & 1 & & & * \\ * & & 1 & & * \\ \vdots & & & \ddots & \vdots \\ * & & & 1 & * \\ * & & & & 1 & * \end{bmatrix}
\quad
\begin{bmatrix} * & 1 & & & \\ & * & 1 & & \\ & & \ddots & \ddots & \\ & & & * & 1 \\ & & & & * & 1 \end{bmatrix}
$$

N_1 $\qquad$ $N_j(2 \leqslant j \leqslant p-1)$ $\qquad$ N_p

Step3:

对 N_p,将本机上的第一行数据中的 $a_{(p-1)k+1}, b_{(p-1)k+1}$ 向上传给 N_{p-1}(对角元 $d_{(p-1)k+1}$ 已经化为 1,不必传,下同)。对 $N_j, 2 \leqslant j \leqslant p-1$,首先接收来自 N_{j+1} 的数据,用接收来的数据消去本机上最后一行数据中的 c_{jk},再将最后一行的对角元化为 1,然后用本机最后一行消去本机第一行数据中的 $c_{(j-1)k+1}$,再将本机第一行数据中的 $a_{(j-1)k+1}$、$b_{(j-1)k+1}$ 向上传给 P_{j-1},最后用本机最后一行消去本机其余各行数据中的 $c_i, (j-1)k+2 \leqslant i \leqslant jk-1$。对 N_1,首先接收 N_2 传来的数据,并用接收来的数据消去本机上最后一行数据中的 c_k,再将最后一行的对角元化为 1。经过 step3 后,各处理机上的系数矩阵形状如下:

$$
\begin{bmatrix} 1 & * & & & \\ & 1 & * & & \\ & & \ddots & \ddots & \\ & & & 1 & * \\ & & & & 1 & 0 \end{bmatrix}
\quad
\begin{bmatrix} * & 1 & & & \\ * & & 1 & & \\ \vdots & & & \ddots & \\ * & & & & 1 \\ * & & & & & 1 & 0 \end{bmatrix}
\quad
\begin{bmatrix} * & 1 & & & \\ & * & 1 & & \\ & & \ddots & \ddots & \\ & & & * & 1 \\ & & & & * & 1 \end{bmatrix}
$$

N_1 $\qquad$ $N_j(2 \leqslant j \leqslant p-1)$ $\qquad$ N_p

Step4:

对 N_1,首先将本机上最后一行数据中的 b_k 向下传送给 N_2,然后自下向上用第 $i+1$ 行消去第 i 行的上对角元 d_i, $1 \leqslant i \leqslant k-1$。对 $N_j, 2 \leqslant j \leqslant p-1$,首先接收 N_{j-1} 传来的数据,用接收来的数据消去本机最后一行数据中的 a_{jk},再将本机最后一行数据中的 b_{jk} 向下传给 N_{j+1},最后用接收来的数据消去本机其余各行数据中的 $a_i, (j-1)k+1 \leqslant i \leqslant jk-1$。对 N_p,首先接收 N_{p-1} 传来的数据,用接收来的数据消去本机上第一行数据中的 a_{n-k+1},然后自下向上用第 $i-1$ 行消去第 i 行数据中的 $a_i, n-k+2 \leqslant i \leqslant n$。

经过 step4 后,各机上的右端项 $b_i(1 \leqslant i \leqslant n)$ 的值就是方程组的解。

2. 算法说明

(1) 将数据平均分配给各台处理机,使各机的负载基本相同,易达到负载平衡。

(2) Step2 中,$N_2 \sim N_{p-1}$ 的向下和向上消元并未使系数矩阵中的非零元减少,但是它使得在 Step3、Step4 中可以不是逐行消元而是跳过中间各行而先消去本机上第一行或最后一行的非对角的非零元,之后立即将消元后的一行数据传给邻机,减少了各处理机之间的等待时间,并且也利用了计算与通信的重叠。在问题规模很大时,各机的等待时间很短,在大部分时间内所有的处理机都在并行地进行本地计算,可获得较高并行度。

(3) 对 N_1 和 N_p,在 Step2 中只进行单边消元,减少了计算量。

(4) 先将对角元消为 1,给邻机传送数据时不必传送对角元,减少了通信量。

3. SPP 算法用于解块三对角方程组

将三对角方程组系数矩阵中的元素换成等阶的方阵，将其右端项的每个元素换为等长度的向量，就成为块三对角方程组的形式。块三对角方程组的解法跟三对角方程组的解法相似，不同之处在于前者中的消去单个元素在后者中变为消去一个矩阵块，这就要将原来的元素相乘/除变为矩阵求逆和矩阵相乘或矩阵乘向量，原来的元素相减变成矩阵相减或向量相减，原来的化对角元为 1 变为化对角块为单位矩阵，另外在后者中邻机间每次传送的数据不是一行而是几行。

4. SPP 算法的性能分析

(1) 存储需求

SPP 算法所需的存储量，对三对角方程组来说为 $4n$(n 为方程组的阶数)，对块三对角方程组来说为 $3ns^2+ns$(块大小为 $s\times s$，方程组的阶数为 $n\times s$)，跟串行的追赶法所需存储量是一样的。

(2) 计算复杂性分析

对三对角方程组，总的乘、除法运算量约为

$$12n-14\frac{n}{p}$$

总的加、减法运算量约为

$$9n-12\frac{n}{p}$$

对块三对角方程组，总的乘、除法运算量约为

$$ns^3+\frac{n}{p}(7p-10)s^3+\frac{n}{p}(4p-4)s^2$$

总的加、减法运算量约为

$$ns^3+\frac{n}{p}(7p-10)s^3+\frac{n}{p}(9p-12)s^2+\frac{n}{p}(4p-4)s$$

与已知最快的串行算法追赶法比较，SPP 算法的运算量要大很多(如对三对角方程组而言，追赶法的乘、除法运算量才为 $5n$，加、减法运算量才为 $3n$)。可见对三对角型方程组的并行直接求解，相对于串行求解来说引入了一定的冗余计算，这是由于三对角型方程组求解的数据相关性强、并行难度大而引起的。各种并行算法都在减少冗余计算量、增加求解过程的并行度、减少通信量和达到负载平衡之间折衷考虑的。对我们后面要介绍的带状和稀疏带状方程组的并行求解也存在同样的问题。

(3) 通信复杂性分析

三对角方程组的 SPP 算法的通信量为 $3(p-1)$，块三对角方程组的 SPP 算法的通信量为 $(p-1)s^2+2(p-1)s$，二者通信建立次数均为 $2(p-1)$。

(4) 性能测试

根据在由 10M 以太网互连的工作站机群(NOWs)上进行的数值实验，得出了如表 8.3、表 8.4 所示的加速比，其中串行计算使用追赶法，块三对角求解中块大小为 3×3，块求逆使用高斯消去法。Sp 为绝对加速比。算式中的执行时间包括文件 I/O 与通信的时间。

表 8.3　三对角方程组的 SPP 算法的测试结果

n	10 000			40 000			160 000			640 000		
p	2	4	6	2	4	6	2	4	6	2	4	6
Sp	1.65	2.38	2.59	1.56	2.09	2.22	1.59	2.33	2.61	1.61	2.16	2.54

表 8.4　块三对角方程组的 SPP 算法的测试结果

n	3 600			14 400			57 600			230 400		
p	2	4	6	2	4	6	2	4	6	2	4	6
Sp	1.88	2.79	3.13	2.02	2.74	2.89	2.06	3.07	3.78	2.01	3.08	3.80

(5) 评述

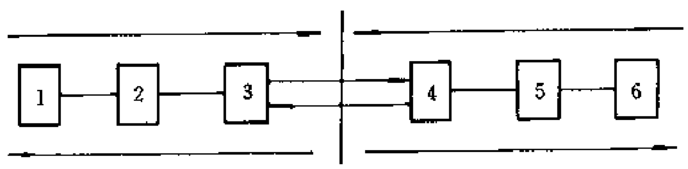
由测试结果可以看出，在使用 2 台处理机时，SPP 算法对三对角和块三对角方程组的加速效率都比较高，这是因为 2 台处理机时的计算量跟串行的追赶法是同样的。而在处理机数多于 2 时，加速比不很高，主要原因是此时 SPP 算法比串行追赶法的计算量大很多，通信启动次数较多、通信时有等待也是部分的原因。

由两张表的对比也可以看出，SPP 算法对块三对角方程组的加速效果比对三对角方程组的要好，原因之一是因为前者的并行粒度比后者的大，原因之二是二者通信启动次数相同，而前者的计算量比后者大很多倍，前者的计算/通信比要高于后者。

二、其他算法介绍

1. DPP 算法

DPP 算法跟 SPP 算法同样使用分治和并行消元，只是在 Step3 时它是从系数矩阵的头和尾同时向中间消元，数据也从两头的处理机（N_1和 N_p）向中间的处理机传送，最后中间的两台处理机互相交换邻接处的一行数据，各自可以求出一个解。之后中间的处理机回代求解，并将边界处一行的数据传给另一个方向相邻的处理机，接收到数据的处理机又可以回代求解，并向另一个方向相邻的处理机传送数据。这一过程继续下去，直到到达两头的处理机。具体算法请看参看迟利华的论文，在此不赘述，仅画出其数据传递顺序示意图如下：



在计算量与通信量上，DPP 算法与 SPP 算法相同。在计算阶段，DPP 算法中部分处理机等待消息的时间比 SPP 算法中的要少，理论上其平均并行度要高一些。

在有些 CFD 计算中，需要由一台处理机收集方程组的解。使用 SPP 算法，并且让数据传送的方向是先由后向前传，后由前向后传，使得编号越小的处理机越早完成计算，用

N_1 按编号顺序接收各个处理机传来的结果，就可以很好地利用计算、文件 I/O 与通信的重叠。此时 SPP 算法也有自己的优势。另外有些情况下，求解完三对角方程组后需要进行一些全局的数据交换操作，SPP 算法也有优势。

2. PPD 算法

在迟利华的论文中还提出了针对严格对角占优三对角方程组的 PPD 算法。其算法思想如下：

对方程组 $Ax = b$，首先将矩阵 A 划分为子矩阵

$$A = \begin{bmatrix} D_1 & C_1 & & & \\ A_2 & D_2 & C_2 & & \\ & \ddots & \ddots & \ddots & \\ & & A_{p-1} & D_{p-1} & C_{p-1} \\ & & & A_p & D_p \end{bmatrix}$$

其中 D_i 为三对角矩阵，$i = 1,2,\cdots,p$。$A_i = a_{(i-1)K+1}e_1e_K^T, i = 2,3,\cdots,p, C_j = C_{jK}e_Ke_1^T, i = 1,2,\cdots,p-1$。$e_1$ 为第一个元素为 1 的 K 维单位向量，e_K 为第 K 个元素为 1 的 K 维单位向量，对右端向量 b 也做相应的划分。

将 A 分解为两个矩阵的乘积

$$A = FT = \begin{bmatrix} D_1 & & & & \\ & D_2 & & & \\ & & \ddots & & \\ & & & D_{p-1} & \\ & & & & D_p \end{bmatrix} \begin{bmatrix} I & H_1 & & & \\ G_2 & I & H_2 & & \\ & \ddots & \ddots & \ddots & \\ & & G_{p-1} & I & H_{p-1} \\ & & & G_p & I \end{bmatrix} \tag{8.76}$$

其中 I 为 $K \times K$ 单位矩阵，$G_i = D_i^{-1}A_i, i = 2,3,\cdots,p, H_j = D_j^{-1}C_j, j = 1,2,\cdots,p-1$。$A_i$ 和 C_j 只有一个非零元素，故 G_i 只有最右边一列元素非零，记为 $g^{(i)}$，H_j 只有最左边一列元素非零，记为 $h^{(j)}$，则

$$D_ig^{(i)} = a_{(i-1)K+1}e_1, \quad i = 1,2,\cdots p \tag{8.77}$$

$$D_jh^j = c_{jK}e_K, \quad j = 1,2,\cdots p \tag{8.78}$$

将 A 分解为(8.76)式后，原方程组的求解就等价于求解

$$Fy = b \tag{8.79}$$

$$Tx = y \tag{8.80}$$

在求解(8.79)的过程中可以同时求出(8.77)和(8.78)。

经证明，$|g_K^{(i)}| \leqslant |g_{K-1}^{(i)}| \leqslant \cdots \leqslant |g_2^{(i)}| \leqslant |g_1^{(i)}| = 1$；对 g_i，若有 $|g_m^{(i)}| = |g_{m-1}^{(i)}|$，则 $|g_m^{(i)}| = |g_{m+1}^{(i)}| = \cdots = |g_K^{(i)}| = 0$。$|h_1^{(j)}| \leqslant |h_2^{(j)}| \leqslant \cdots \leqslant |g_{K-1}^{(j)}| = 1$；对 h_j，若有 $|h_m^{(j)}| = |h_{m-1}^{(j)}|$，则 $|h_m^{(j)}| = |h_{m-1}^{(j)}| = \cdots = |h_1^{(j)}| = 0$。故若 K 足够大，对任意正数 δ_1，总存在正整数 K_{i1} 和 K_{i2}，使得 $|g_k^{(i)}| < \delta_1, k = K_{i1},\cdots,K$；且 $|h_k^{(i)}| < \delta_1, k = 1,\cdots,K-K_{i2}$ 成立。

设 x 为方程组(8.80)的精确解，给 T 一个扰动 ΔT，

$$\Delta T = \begin{bmatrix} \Delta T_1 & & & \\ & \Delta T_2 & & \\ & & \ddots & \\ & & & \Delta T_p \end{bmatrix} \tag{8.81}$$

其中

$$\Delta T_i = \begin{bmatrix} 0 & 0 & \cdots & 0 & -h_1^{(i)} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -h_{K-K_{i2}}^{(i)} \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ -g_{K_{i1}}^{(i)} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ -g_K^{(i)} & 0 & \cdots & 0 & 0 \end{bmatrix}, i = 1,2,\cdots,p \tag{8.82}$$

K_{i1}和K_{i2}为正整数，使$\|T^{-1}\|_\infty \|\Delta T\|_\infty < 1$，令$\delta = \max\limits_i(\max(|g_{K_{i1}}^{(i)}|, |h_{K-K_{i2}}^{(i)}|))$，解方程组$(T+\Delta T)(x+\Delta x) = y$，给定误差界$\varepsilon$，则当$\delta$满足

$$\delta < \frac{1}{2} \cdot \frac{\varepsilon}{(1+\varepsilon)} \cdot \frac{1}{\|T^{-1}\|_\infty} \tag{8.83}$$

时，必有

$$\frac{\|\Delta x\|_\infty}{\|x\|_\infty} < \varepsilon \tag{8.84}$$

成立。迟利华还证明了当$q = \max(\max\limits_i(K_{i1}), \max\limits_i(K_{i2}))$满足一定的条件时，方程组$(T+\Delta T)(x+\Delta x) = y$的解满足不等式(8.84)。

根据上述推导，给定误差界ε，如果方程组的阶数n足够大，可以通过求解$Fy = b$和$(T+\Delta T)x = y$来求出三对角方程组的满足误差界ε的近似解。PPD 算法的过程如下：

Step1：

将$D_i, a_{(i-1)K+1}, c_{iK}, b_{(i-1)K+1}, \cdots, b_{iK}$分配到第$i$台处理机上，$i = 1,2,\cdots p$。

Step2：

各台处理机同时进行消元，第i台处理机消去D_i的对角线以下元素并将对角元素化为1，但对第2至第p台处理机的对角线下新引入的非零元，当其某个值小于开关值δ后，下面新引入的非零元都近似看做零元素，向下消元时不再计算新引入的非零元。消元后处理机i将c_{iK}与b_{iK}传送给处理机$i+1$，$i = 1,2,\cdots p-1$。

Step3：

各台处理机同时进行消元，第i台处理机消去D_i的对角线以上元素，但对第1至第$p-1$台处理机的对角线以上新引入的非零元，当其某个值小于开关值δ后，后面新引入的非零元都近似看做零元素，向上消元时不再计算新引入的非零元。

经过 Step2 和 Step3 之后，方程组的系数矩阵被化为如下形式所示(只画出中间相邻的两台处理机上的系数矩阵的形式，并忽略了全为零的块)：

$$
\underset{N_i}{\begin{bmatrix}
g_1^{(i)} & 1 & & & & & 0 \\
g_2^{(i)} & & 1 & & & & \vdots \\
\vdots & & & \ddots & & & 0 \\
g_{K_1}^{(i)} & & & & \ddots & & h_{K-K_{i2}-1}^{(i)} \\
0 & & & & & \ddots & \vdots \\
\vdots & & & & & 1 & h_{K-1}^{(i)} \\
0 & & & & & & 1 \quad h_K^{(i)}
\end{bmatrix}}
\qquad
\underset{N_{i+1}}{\begin{bmatrix}
g_1^{(i+1)} & 1 & & & & & 0 \\
g_2^{(i+1)} & & 1 & & & & \vdots \\
\vdots & & & \ddots & & & 0 \\
g_{K_1}^{(i+1)} & & & & \ddots & & h_{K-K_{i2}-1}^{(i+1)} \\
0 & & & & & \ddots & \vdots \\
\vdots & & & & & 1 & h_{K-1}^{(i+1)} \\
0 & & & & & & 1 \quad h_K^{(i+1)}
\end{bmatrix}}
$$

Step4：

第 $i+1$ 台处理机接收第 i 台处理机传来的数据，计算

$$
\begin{bmatrix} 1 & c_{iK} \\ g_1^{(i+1)} & 1 \end{bmatrix} \begin{bmatrix} x_{iK} \\ x_{iK+1} \end{bmatrix} = \begin{bmatrix} b_{iK} \\ b_{iK+1} \end{bmatrix} \tag{8.85}
$$

并将 x_{iK+1} 的值传送给第 i 台处理机，$i=1,2,\cdots,p-1$。

Step5：

第 i 台处理机利用 $x_{(i-1)K+1}$ 和 x_{iK+1} 的值回代求出 $x_{(i-1)K+2},\cdots,x_{iK}$。

PPD 算法有两个好处，一是 Step2 和 Step3 中向下或向上消元时，当计算出的新非零元小于开关值 δ 后，消元中就不必考虑新引入的非零元的值，这就节省了大量的计算开销；二是 Step3 向下消元后，第 i 台处理机可以立即将 c_{iK} 与 b_{iK} 传给处理机 $i+1$，而 Step4 中第 i 台处理机解完一个二阶方程组后，可以立即将 x_{iK+1} 的值传递给第 $i+1$ 台处理机，不像原来的分裂法如 SPP 和 DPP 一样需要一个传递等待的过程，减少了通信等待的时间。根据迟利华的测试结果，当方程组的阶很大时，并行计算的加速比接近处理机台数。对严格对角占优三对角方程组，这是一个很优秀的算法。

8.6.2 带状五对角和块带状五对角方程组的直接并行计算

对一维问题进行五点中心差分离散，可以得到带状五对角或者块五对角方程组。其形式如下图，其中对带状块五对角方程组而言，系数矩阵中的元素为等大小的块，未知量和右端项元素为等大小的向量。

$$
\begin{bmatrix}
d_1 & e_1 & f_1 & & & & \\
b_2 & d_2 & e_2 & f_2 & & & \\
a_3 & b_3 & d_3 & e_3 & f_3 & & \\
 & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & a_{n-2} & b_{n-2} & d_{n-2} & e_{n-2} & e_{n-2} \\
 & & & a_{n-1} & b_{n-1} & d_{n-1} & e_{n-1} \\
 & & & & a_n & b_n & d_n
\end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix}
=
\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{n-2} \\ b_{n-1} \\ b_n \end{bmatrix}
$$

对五对角方程组或者块五对角方程组的直接解法，主要有两种思路，一是推广三对角型方程组的分裂法到五对角型方程组的情况，二是直接将其系数矩阵分块，形成块三对角形式的方程组，再使用块三对角方程组的一些并行算法进行求解。下面对这两类方法进行介绍。

一、带状五对角方程组的并行分裂法

根据五对角矩阵与三对角矩阵形式相似的特点，也可以使用各种分裂法来对五对角方程组进行并行求解。这里使用跟三对角方程组的 DPP 算法相似的消元过程，通过文字描述加图示的方法来说明这种算法(或者我们可以称之为五对角方程组的 DPP 算法)。

1. 算法概述

数据按行依顺序平均分配到各处理机上，存储五条对角线上的元素和右端项，消元过程中引入的新的非零元就存储在刚被消去的元素的位置。且假定处理机台数为 p，p 能被 2 整除，给各处理机编号为 $N_1, N_2, \cdots, N_p$。

并行消元的过程是这样的：首先，各处理机并行地向下和向上消元，并将对角元化为 1，直到各机上的系数矩阵消为如下的形式(图中忽略了两侧元素全为零的块，以下同)：

$$
\begin{bmatrix}
1 & & & & * & * \\
& 1 & & & * & * \\
& & \ddots & & \vdots & \vdots \\
& & & 1 & * & * \\
& & & & 1 \;\; * & *
\end{bmatrix}
\quad
\begin{bmatrix}
* & * & 1 & & & & * & * \\
* & * & & 1 & & & * & * \\
\vdots & \vdots & & & \ddots & & \vdots & \vdots \\
* & * & & & & 1 & * & * \\
* & * & & & & & 1 \;\; * & *
\end{bmatrix}
\quad
\begin{bmatrix}
* & * & 1 & & & \\
* & * & & 1 & & \\
\vdots & \vdots & & & \ddots & \\
* & * & & & & 1 \\
* & * & & & & & 1
\end{bmatrix}
$$

N_1 $\qquad\qquad$ $N_i(1 < i < p)$ $\qquad\qquad$ N_p

注意到此时第一与第二台处理机交界处的矩阵元素的形式，接下来的一步，N_1 将本机上最后两行的非零元素下传给 N_2；N_2 接收到第一台处理机传来的数据后，可以将本机上最前两行元素中处在对角线左边的两个非零元消去，而且将新引入的非零元也消去，然后利用接收来的两行数据和本机上最前两行数据，可以将本机上最后两行元素中处在主对角线左边的两个非零元消去，而且将新引入的非零元也消去。N_p 与 N_{p-1} 之间也可以进行类似的消元，只是方向相反。然后 N_2 与 N_3 之间、N_{p-1} 与 N_{p-2} 之间又可以进行类似的操作，而且 N_2 在将本机上最后两行中的非零元传给 N_3 后，可以接着将本机上剩下的处在主对角线左边的非零元全部消去，N_{p-1} 也同样。这样从两个方向同时向中间并行消元，只是到 $N_{p/2}$ 不再向下传数据，$N_{p/2+1}$ 不再向上传数据。直到 $N_{p/2}$ 将其对角线左边的非零元全部消去，$N_{p/2+1}$ 将其对角线右边的非零元全部消去。此时各处理机上的系数矩阵变为如下形式(N_1 与 N_p 的系数矩阵跟上图比未变，故未画出)：

$$
\begin{bmatrix}
1 & & & & * & * \\
& 1 & & & * & * \\
& & \ddots & & \vdots & \vdots \\
& & & 1 & * & * \\
& & & & 1 \;\; * & *
\end{bmatrix}
\qquad
\begin{bmatrix}
* & * & 1 & & & \\
* & * & & 1 & & \\
\vdots & \vdots & & & \ddots & \\
* & * & & & & 1 \\
* & * & & & & & 1
\end{bmatrix}
$$

$N_i(2 \leqslant i \leqslant p/2)$ $\qquad\qquad$ $N_j(p/2+1 \leqslant j \leqslant p-1)$

再下一步，$N_{p/2}$ 和 $N_{p/2+1}$ 互相交换邻接处的两行数据中的非零元，$N_{p/2}$ 用接收来的数据与本机的最后两行数据构成一个可解的缩减方程，解此缩减方程，再将其解代入本机上最前两行方程，可以消去最前两行数据中主对角线右边的非零元，之后将最前两行数据中的非零元向前传给 $N_{p/2-1}$，再消去本机上主对角线右边余下的非零元。在 $N_{p/2+1}$ 与 $N_{p/2+2}$ 之间也可以进行类似的操作。这样由中间向两头传递，直到 N_1 与 N_p。各处理机并行消元，

直到各机上将所有的非对角元都消为零元素。

上述步骤中，系数矩阵消元时，相应的右端项也随之变化。消元结束时，各处理机上右端项的集合就是原方程组的解。

对块五对角方程组也可使用同样的并行消元法，只是上述步骤中对单个元素的操作要变为对矩阵或者向量的操作。

二、带状五对角方程组的块三对角并行求解法

将方程组系数矩阵沿对角线及其两侧分成 2×2 的块，则系数矩阵就可看成是块三对角阵的形式，如下图：

$$\begin{bmatrix} \begin{bmatrix} \ast & \ast \\ \ast & \ast \end{bmatrix} & \begin{bmatrix} \ast & \ast \\ \ast & \ast \end{bmatrix} & & & \\ \begin{bmatrix} \ast & \ast \\ & \ast \end{bmatrix} & \begin{bmatrix} \ast & \ast \\ \ast & \ast \end{bmatrix} & \begin{bmatrix} \ast & \\ \ast & \ast \end{bmatrix} & & \\ & \ddots & \ddots & \ddots & \\ & & \begin{bmatrix} \ast & \ast \\ & \ast \end{bmatrix} & \begin{bmatrix} \ast & \ast \\ \ast & \ast \end{bmatrix} & \begin{bmatrix} \ast & \\ \ast & \ast \end{bmatrix} \\ & & & \begin{bmatrix} \ast & \ast \\ & \ast \end{bmatrix} & \begin{bmatrix} \ast & \ast \\ \ast & \ast \end{bmatrix} \end{bmatrix}$$

使用上面所讨论的三对角型方程组的并行求解方法如 SPP 算法或 DPP 算法，可以很容易地进行求解。鉴于上面已经详细讨论过这些解法，此处就不再重复。

对块五对角方程组也可使用同样的分块和求解方法，无非是块更大而已。

有些计算中为了追求高精度，需要对一维问题进行七点中心差分离散，这就得到七对角方程组或者块七对角方程组。对七对角型方程组的直接求解，跟五对角型方程组类似，同样可以推广到使用并行分裂法求解，只是交换的数据量更大；或者直接将其系数矩阵分块（分块大小为 3×3），形成块三对角形式的方程组，然后再使用块三对角方程组的一些并行算法进行求解。具体算法请读者自行推导，在此不赘述。

8.6.3 稀疏带状方程组的并行计算

对于稀疏带状方程组如稀疏带状五、七对角和稀疏带状块五、七对角方程组，除了可以用上面介绍的迭代法并行求解外，还可以使用直接法并行求解。

因为稀疏矩阵跟一个邻接图相对应，一些研究中采用先对邻接图的节点进行有效的排序如嵌套排序、极小度排序、扇入排序等，排序之后再采用对相应的新矩阵进行分解的方法来求解方程组，分解时通常并不能保证原矩阵中元素为零的位置上其分解因子的元素也为零，即会引入了一些填入元（fill-in）。基于排序的算法可以有效地减少存储需求与计算量，但处理起来比较复杂，而且数据相关性也比较强，并行难度大。我们手头尚无这方面的较好的算法资料。

因为块三对角方程组已经有较成熟的并行解法，有些研究中把稀疏带状五、七对角型方程组归结为块三对角方程组，使用解块三对角方程组的方法对其进行并行求解。这些算

法在简单有效的同时，但计算量很大。

总之，大型稀疏带状线性方程组的并行求解是个难题。特别对并行直接求解来说，由于数据相关性强，求解中如何达到较高的并行度，同时尽量减少冗余计算量和通信开销，是很困难的问题。

我们也提出了一种跟 Gauss 消去法类似的算法(PPE 算法)，在一定程度上发掘了问题的并行性。下面将介绍分块后并行分裂的算法和 PPE 算法。

一、稀疏带状五对角和块五对方程组的并行计算

对二维问题进行强隐格式的五点中心差分离散，就可以得到稀疏带状五对角(非耦合求解)或者稀疏带状块五对角方程组(耦合求解)。设在两个方向上布置的网格点数分别为 $jm+1$、$im+1$，则方程组的系数矩阵如下图所示：

$$\begin{bmatrix}
d_1 & g_1 & & & h_1 & & & & \\
e_2 & d_2 & g_2 & & & h_2 & & & \\
 & e_3 & d_3 & g_3 & & & \ddots & & \\
 & & & & & & & h_{im\times jm-jm-1} & \\
f_{jm+1} & & & \ddots & \ddots & \ddots & & & h_{im\times jm-jm} \\
 & f_{jm+2} & & & \ddots & \ddots & \ddots & & \\
 & & \ddots & & & & & & \\
 & & & f_{im\times jm-1} & & & e_{im\times jm-1} & d_{im\times jm-1} & g_{im\times jm-1} \\
 & & & & f_{im\times jm} & & & e_{im\times jm} & d_{im\times jm}
\end{bmatrix}$$

其中次对角线 f 到 e、g 到 h 之间的带宽为 $jm-2$，这一带宽一般比较大。根据五点中心差分离散格式，对于 e、g 两条次对角线上的元素，每隔 $jm-2$ 个元素就有一个零元。在本文中称每一条平行于对角线的直线上的所有元素的集合为一层。

1. *稀疏带状五对角型方程组的块三对角并行求解法*

我们也可以将上述系数矩阵划分为块三对角阵，然后使用块三对角的并行算法来对其进行求解。由于 e、g 两条次对角线每隔 $jm-2$ 个元素就有一个零元，因此正好可以将系数矩阵划分成块大小为 $jm\times jm$ 的块三对角阵。而且，在此块三对角方程组中的每个块不是对角阵就是三对角阵。

在使用分裂法如 SPP 分裂法对这种块三对角方程组进行并行求解中，在 Step2 进行块消去和化对角块为单位阵的开始阶段，可以利用初始块为对角阵或者三对角阵的特点来减少计算量。但是，消去完几个块后，新引入的块就几乎变成了稠密阵了。而 Step2 结束时，所有新引入的位于主对角块两边的块都是稠密的了。所以我们还必须将这些块当做稠密阵来存储，即相当于存储从次对角线 f 到次对角线 h 的整个带状部分(包括开始的非零元)。在计算中，为了简便，也把所有的块都当做稠密阵来处理。

对稀疏带状块五对角方程组也可以用同样的算法。

这种算法在简单化的同时，也带来存储量需求巨大，计算量高的问题。特别是对这么大的块要进行求逆、相乘等操作，不仅计算量大，而且需要大量的数据移动，这是非常费时的。

下面介绍另外一种思路的并行算法，我们暂且称其为流水线并行消去法 PPE

(Pipeline Parallel Elemination)算法。

2. 稀疏带状五对角方程组的 PPE 算法

考虑到次对角线 f 到到 h 之间的带宽较大(总带宽为 $2\times jm+1$)的特点,如果将数据按行分配到各处理机上,不同的处理机同时消去本机上不同层上的非零元素,则可以获得较高的并行度。当然,各处理机对同一层的元素进行消元时有一定的顺序,这就需要数据传递和等待。这就是 PPE 算法的核心思想。

(1) 算法描述

假设有 p 台处理机,编号为 $N_1, N_2, \cdots, N_p$,不失一般性,设 $p\times k=im\times jm$,并规定 $k>jm$。数据按行平均分配到各台处理机上,即每台处理机分配到系数矩阵中的 k 行和右端向量中的 k 个分量。存储从对角线 f 到对角线 h 的整个带状部分,也存储带状部分中间初始为零的元素,因为在消元过程中会在这些位置引入非零元。另外还要存储右端向量。算法步骤如下:

Step1:

对 N_p:

    首先将第 $k-jm$ 到第 $k-2$ 行数据乘以一定的系数后加到第 $k-jm+1$ 行上
    for *layer* from 1 to $jm-1$ do
        用第 $k-layer$ 行消去第 $k-jm$ 行的当前最右一个非零元
        依次用第 $k-jm$ 行到第 2 行消去其上一行的当前最右一个非零元
        将本机第一行数据上传给 N_p-1
    end for
    依次用第 k 行到第 2 行消去其上一行的当前最右一个非零元
    将本机第一行数据上传给 N_p-1

对 $N_i(i=2,\cdots,m-1)$:

    for *layer* from 1 to *jm* do
        等待并接收来自 N_i-1 的数据
        用接收来的数据消去本机最后一行数据中的当前最右一个非零元
        依次用第 k 行到第 2 行消去其上一行的当前最右一个非零元
        将本机第一行数据上传给 N_i-1
    end for

对 N_1:

    for *layer* from 1 to *jm* do
        等待并接收来自 N_2 的数据
        用接收来的数据消去本机最后一行数据中的当前最右一个非零元
        依次用第 k 行到第 2 行消去其上一行的当前最右一个非零元
    end for

经过 Step1 后，系数矩阵主对角线右边的非零元完全消去。

Step2:

对 N_1：

for *layer* from 1 to *jm* − 1 do

用第 *layer* + 1 行消去第 *layer* + 2 到 *jm* + 1 行的当前最左一个非零元

依次用第 *jm* + 1 行到第 *k* − 1 行消去其下一行当前最左一个非零元

将本机最后一行数据向下传给 N_2

end for

依次用第 1 行到第 *k* − 1 行消去其下一行的当前最左一个非零元

将本机最后一行数据向下传给 N_2

对 $N_i(i = 2,\cdots,m-1)$：

for *layer* from 1 to *jm* do

等待并接收 N_{i-1} 传来的一行数据

用接收来的数据消去本机第一行的当前最左边的非零元

依次用第 1 行到第 *k* − 1 行消去其下一行的当前最左一个非零元

将本机最后一行数据向下传给 N_{i+1}

end for

对 N_p：

for *layer* from 1 to *jm* do

等待并接收 N_p − 1 传来的一行数据

用接收来的数据消去本机第一行的当前最左边的非零元

依次用第 1 行到第 *k* − 1 行消去其下一行的当前最左一个非零元

end for

Step2 跟 Step1 类似，只是消元方向为由上向下。它将系数矩阵主对角线左边的非零元完全消去。

Step3:

所有的处理机并行地将本机上的系数矩阵对角元化为 1。

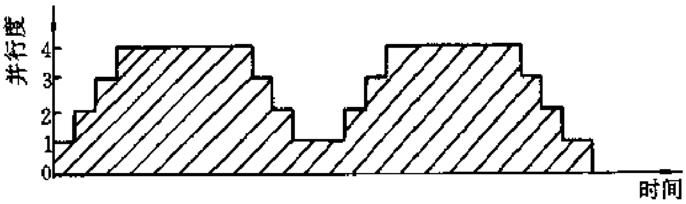
在对系数矩阵进行消元时，方程组的右端项也相应变化。经过 Step3 后，右端项即为所求的方程组之解。注意上述算法描述中所用的“第 x 行”是相对于本机而言的。

(2) 理论分析及结果测试

• PPE 算法的并行模式

PPE 算法的并行模式跟微处理器中指令的流水线并行执行相似，算法中的每台处理机就相当于流水线中的一站(stage)，系数矩阵中的一层元素就相当于流水线中执行的一条指令。我们这里假设 $p < jm$。在 Step1 中，N_p 首先消去完本机上系数矩阵右边的第一层，将本机上第一行数据传给 N_{p-1}，于是 N_{p-1} 可以开始消本机上系数矩阵的右边第一层，此时 N_p 开始消右边第二层，N_{p-1} 消完右边第一层又向 N_{p-1} 传本机上第一行数据，

……，如此传递下去，到某一时刻开始所有的处理机都在进行消元操作，此时并行度为 p，保持并行度为 p 一段时间后，N_p 首先消完本机上系数矩阵的主对角线右边的所有非零元，停止上消，接着 N_{p-1} 到 N_1 依次停止上消。Step2 跟 Step1 类似，只不过是 N_1 首先开始向下消元并最先将主对角线左边的非零元消完。这跟流水线执行时有一个启动(start-up)、满流(full)和清空(clean-up)的过程是很相像的。程序执行中的并行度与时间的关系如下(以 $p = 4$ 为例)：



不难得出，系数矩阵次的次对角线 e 到次对角线 h 的带宽越大(即 jm 越大)，则程序执行期间的平均并行度就越大，并行效率会越高。

· PPE 算法的存储需求

由于需要存储次对角线 e 到次对角线 h 中间的所有带状部分，加上右端项，共需 $im \times jm \times (2 \times jm + 2)$ 个存储单元。

· PPE 算法计算与通信开销

估算了 PPE 算法和分块算法(解块三对角方程组用单向并行分裂法 SPP)的计算量和通信开销，其中计算量以浮点乘除法的总量(浮点加、减法的总量跟乘除法的相近)为标准，并略去了所有的小量，结果如表 8.5。

表 8.5 PPE 算法与分块算法的计算与通信开销比较

算法	计算量	通信量	通信启动次数
分块算法	$(4 - 4/p)im \cdot jm^2 + (8 - 10/p)im \cdot jm^3$	$2(p-1)jm(jm+1)$	$2(p-1)$
PPE 算法	$2im \cdot jm^3$	$2(p-1)jm^2$	$2(p-1)jm$

由表 8.5 可见，PPE 算法的计算量约是分块算法计算量的 1/6。在通信量方面，PPE 算法与分块算法的接近，但 PPE 算法的通信启动次数比分块算法的要多。二者的计算量都约与 im 成正比，与 jm 的立方成正比。

· 测试结果

用 PPE 算法和分块算法编制并行程序，并用变形的高斯消去法(不分块，消元只消到带状边界，其计算量经估算约为 $im * jm^3$)编制串行计算程序，在工作站机群上进行测试，取得数据如表 8.6、表 8.7，其中加速比为解相同规模问题时，并行计算的墙上时间跟串行算法计算的墙上时间之比。对串行程序，$im = 72, jm = 20$ 时的墙上时间为 0.2119s；$im = 72, jm = 40$ 时的墙上时间为 2.5428s.

表 8.6　测试结果一 ($im = 72, jm = 20$)

处理机数		2	4	6
分块算法	墙上时间	1.1599	0.8663	0.3624
	加速比	0.18	0.25	0.58
PPE 算法	墙上时间	0.1742	0.1423	0.1195
	加速比	1.22	1.49	1.77

表 8.7　测试结果二 ($im = 72, jm = 40$)

处理机数		2	4	6
分块算法	墙上时间	11.6863	5.8021	5.6072
	加速比	0.22	0.47	0.49
PPE 算法	墙上时间	1.9339	0.9445	0.7078
	加速比	1.31	2.69	3.59

· 评述

由表 8.6、表 8.7 可以看出，分块算法的墙上时间很长，而 PPE 算法的墙上时间只是分块算法的几分之一，说明 PPE 算法的计算效率确实比基于分块的算法高。当带宽扩大时，程序的执行时间迅速增长，但是 PPE 算法的加速比也有较大的增长，跟前面的分析是吻合的。分块算法的加速比还不到 1，是因为它的计算量是串行算法的很多倍的缘故。

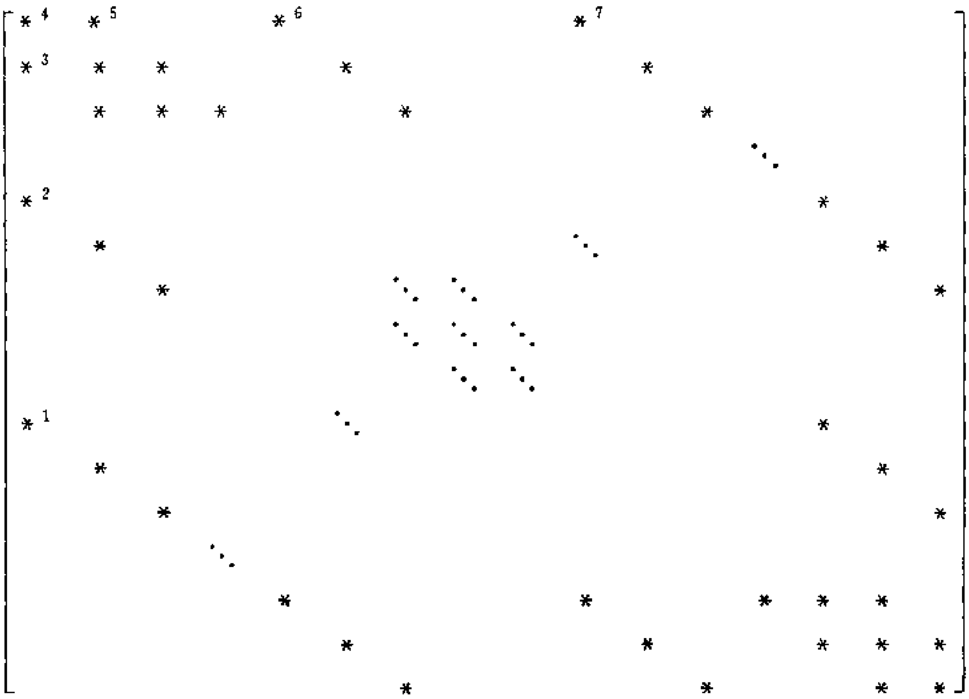
另外也应注意到，PPE 算法一方面节省了总计算量，另一方面却增大了通信启动次数，通信开销的影响较为显著。PPE 算法在测试结果一中的加速比较低，这主要因为其带宽窄，im 也较小，导致计算计算/通信比较小，并行效率不高。

上面两种算法都不是很理想，除了存储需求巨大之外，还存在计算量巨大的问题，特别是基于分块的并行分裂法的计算量很大。另一方面，PPE 算法的实质是串行的高斯消去法的变形形式，求解中几乎不能选主元，而分块算法虽然在块求逆中可以选主元，但其实也很有限。这就带来求解过程的稳定性问题，在问题规模很大时，舍入误差可能会严重积累，使得计算出来的结果偏差很大。控制误差跟使求解过程高度并行是一对矛盾的问题，难以找到很好的解决办法，这也是很多应用中都尽量避免稀疏带状方程组的并行直接求解的原因。排序之后再并行分解与消去的方法在稳定性方面要好一些，计算量也得到减少，但我们这里尚没有该方面并行算法的资料。有兴趣的读者可以在这方面进行深入研究。

上述两种算法可以推广到稀疏块带状五对角方程组的并行求解上。

二、稀疏带状七对角和块七对角方程组的并行求解

对三维问题进行强隐格式的七点中心差分离散，就可以得到稀疏带状七对角(非耦合求解)或者稀疏块带状七对角方程组(耦合求解)。稀疏带状七对角方程组的系数矩阵如下所示：



设在三个方向上布置的网格点数分别为 $im+1$、$jm+1$、$km+1$，则次对角线 1 与次对角线 2 之间的间隔为 $km-2$，次对角线 2 与次对角线 3 之间的间隔为 $jm\times km-km-1$。

对于稀疏带状七对角或稀疏带状块七对角方程组的并行求解，可以使用跟稀疏带状五对角或稀疏块带状五对角方程组求解的两种方法类似的策略，这里不再详述。但是由于这两种算法用于稀疏带状七对角方程组时的存储开销巨大，计算量也很大，而且还存在稳定性问题，故对大型的问题是不适用的。对大型的三维问题，如果直接离散导出稀疏带状七对角型方程组的话，我们还是倾向于使用迭代法来解。

参 考 文 献

[1] Mitchell A R, etal. The Finite Difference Method in Partial Differential Equations. John Wiley & Sons, 1980

[2] Patankar S V 著. 郭宽良译. 传热与流动的数值方法. 合肥:安徽科学出版社,1984

[3] Connor J J 著. 吴望一译. 流体流动的有限元法. 北京:科学出版社,1981

[4] 马铁犹. 计算流体力学. 北京:北京航空学院出版社,1988

[5] 傅德薰. 流体力学数值模拟. 北京:国防工业出版社,1993

[6] 朱自强. 应用计算流体力学. 北京:北京航空航天大学出版社,1998

[7] Anderson J D Jr. Computational Fluid Dynamics. McGraw-Hill Inc., 1995

[8] Fletcher C A J. Computational Techniques for Fluid Dynamics (Vol. I, Vol. II). Springer-Verlag, 1988

[9] Wendt J F. Computational Fluid Dynamics. Springer-Verlag, 1992

[10] 陆金甫,关治. 偏微分方程数值解法. 北京:清华大学出版社,1987

[11] 刘兴隆,郑群. 计算流体力学. 哈尔滨:哈尔滨工业大学出版社,1998

[12] 苏铭德,黄素逸. 计算流体力学基础. 北京:清华大学出版社,1997

[13] 李松波. 耗散守恒格式理论. 北京:高等教育出版社,1997

[14] 徐华舫. 空气动力学基础(上、下册). 北京:国防工业出版社,1980

[15] 吴建民. 高等空气动力学. 北京:北京航空航天大学出版社,1992

[16] 童秉纲等. 气体动力学. 北京:高等教育出版社,1990

[17] 赵学端等. 粘性流体力学. 北京:机械工业出版社,1993

[18] 张涵信. 发展计算空气动力学应该解决的几个重要问题. 空气动力学研究文集,1994(4)

[19] Rai M M. A Relaxation Approach to Patched Grid Calculation with the Euler Equations. AIAA paper, 85—0195

[20] Yoon S, Jameson A. An LU-SSOR Scheme for the Euler and NS Equations. AIAA J., 1987, 26 (9)

[21] Rieger H, Jameson A. Solution of Steady Three-Dimensional Compressible Euler and NS Equations by Implicit LU Scheme. AIAA paper, 88—0619

[22] Yeneriz M A, etal. Comparison of Calculation and Experiment for a Lateral Jet from a Hypersonic Bianic Vehicle. AIAA paper 89—2548

[23] 梁剑寒,王承尧. 超燃发动机燃烧三维化学反应流场的数值模拟. 航空学报,1998(3)

[24] Jameson A. Analysis and Design of Numerical Schemes for Gas Dynamic Artificial Diffusion, Upward Biasing, Limiters and Their Effect on Accuracy and Multigrid Convergence. N95-11875

[25] Beam R W, Warming R F. An Implicit Finite Difference Algorithm for Hyperbolic Systems in Conservation Form J. comp. phys, 1976(22)

[26] MacCormack R W. The Effect of Viscosity in Hyper-Velocity Impact Cratering. AIAA paper, 69—0354

[27] Pulliam T H, Chaussee D S. A Diagonal Form of an Implicit AF Algorithm. J. comp. phys., 1989(39)

[28] Harten A. High Resolution Schemes for Hyperbolic Conservation Laws. J. comp. phys, 1983(49)

[29] Chakravanthy S R, Osher S. A New Class of High Accuracy TVD Scheme for Hyperbolic Conservation Laws. AIAA paper, 85—0363
[30] Yee H C. Upwind and Symmetric Shock—Capturing Schemes. NASA, TM—89464 1987
[31] 张涵信.无波动、无自由参数的耗散格式.空气动力学报, 1988,6(2)
[32] 邓小刚,张涵信.NND 格式的推广及其在粘性计算中的应用.空气动力学报, 1994,12(2)
[33] Jameson A. Success and Challenges in Computation Aerodynamics. AIAA paper, 87—1184
[34] Chakravathy S R. High Resolution Formalations for the NS Equations. N89—17824
[35] 张涵信.差分计算中激波上下游解出现波动的探讨.空气动力学报,1984(1)
[36] 张涵信.求解气动方程高精度格式及其相关的问题.第七届全国计算流体力学会议论集,1994
[37] Li H, Wang C Y. Multiblock/Multigrid Implicit TVD Solution of 3D Viscous Hypersonic Flow. AIAA paper, 96—0314
[38] 王承尧,李桦.用多重网格 TVD 法求解大攻角高超音速三维粘性流场.空气动力研究论文集, 1994(4)
[39] MacCormack R W. A Perspective on a Quarter Century of CFD Research. AIAA paper, 93—3291
[40] Swanson R C, Twrkel E. Aspects of a High-Resolution Scheme for the NS Equations. AIAA paper, 93—3372
[41] Swanson R C, Twrkel E. Artificial Dissipation and Central Difference Schemes for Euler and NS Equations. AIAA paper, 87—1107
[42] Thompson J F. Grid Generation Techniques in CFD. AIAA J, 1984,22
[43] Steger T L, Warming R F. Flux Vector Splitting of the Invisid Gasdynamics Equations with Application To Finite Difference Method. J. comp, phys, 1981,40
[44] Harten A. ENO Schemes with Subcell Resolution. J. comp. phys., 1989,83
[45] Liou M S. A Flux Splitting Scheme with High Resolution and Robustness for Discontinuation. N 94—36646
[46] Godunov, S K. A Finite Difference Method for Numerical Computation of Discontinuance Solutions of the Equations of Fluid Dynamics. Soviet Math. sb., 1959(47)
[47] Swanson R C, Turkel E. On Central Difference and Upwind Schemes. J. comp, phys., 1992 (101)
[48] Van Leer B. Flux Vector Splitting for Euler Equations. Lecture Notes in phys, 1982(172)
[49] Swely P K. High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. SIAM Journal Num. Anal., 1984(21)
[50] Zhang H X, Zhuang F G. NND Schemes and Their Applications to Numerieal Simulation of Two and Three Dimensional Flows. Advances in Applied Mechanics, 1992(29)
[51] Yee H C. Construction of Explicit and Implicit Symmetric TVD Schemes and Their Application. J. comp. phys., 1987(68)
[52] Thompson J F, etal. Aspects of Numerical Grid Generation: Current and Art. AIAA paper, 93—3539
[53] Benek J A, etal. A 3—D Chimera Grid Embedding Technique. AIAA paper, 85—1523
[54] Wilcox D C, Rubesin M W. Progress in Turbulence Modeling for Complex Flowfields Including Effects of Compressibility. NASA, TP—1517, 1980
[55] 柳军,瞿章华.再入飞行器高超声速三维粘性流场的数值模拟.国防科技大学学报,1996(2)
[56] 石于中,陈伟芳,任兵,吴其芬.稀薄流热化学非平衡效应的 DSMC 研究.空气动力学学报,1997

(2)

[57] Grasso F. Marimi M. Solution of Hypersonic Viscous Flows with Total Variation Diminishing Multigrid Techniques. Computers & Fluids, 1995(5)

[58] Rumsey C L, etal. Efficiency and Accuracy of Time-Accurate Turbulent N. S. Computations. Computers & Fluids. 1996(2)

[59] Choi Y H, Merkle C L. The Application of Preconditioning in Viscous Flows. J. comp. phys. 1993(105)

[60] Kim C A. Jameron A. Flux Limited Dissipation Schemes for High Speed Unsteady Flows. AIAA paper, 95－1738

[61] Batten P, etal. A Time-Splitting Approach to Solving the NS Equation. Computers & Fluids. 1996 (4)

[62] Govindaw T R, Lakshminarayane B. Computation of Viscous Internal Flows－A Space-Marching Method. Computers & Fluids, 1988(1)

[63] Wang Z J, Richards B E. High Resolution Schemes for Steady Flow Computation. J. comp. phys, 1990(87)

[64] Liou M S. Progress Toward An Improved CFD Method: AUSM$^+$. AIAA paper, 95－1701

[65] Radespiel R, Krdl N. Accurate Flux Vector Splitting for Shocks and Shear Layers. J. comp. phys, 1995(121)

[66] Gaitonde D, Shang J S. Accuracy of Flux-Split Algorithms in High-Speed Viscous Flows. AIAA J., 1993(7)

[67] Godfrey A G, Michell C R. Parctical Aspects of Spatially High-Order Accurate Methods. AIAA J., 1993(9)

[68] Van Leer B. Flux-Vector Splitting for the 1990s. N91－21073

[69] Roe P L. Approximate Rieman Slovers, Parameter Vectors and Difference Sohemes. J. comp. phys, 1981(43)

[70] Andevron W K, etal. A Comparision of Finite Volume Flux Vector Splitting for Euler Equations. AIAA paper, 85－0122

[71] Van Leer B. Towards The Ultimate Conservation Difference Scheme Ⅰ. Monotonicity and Conversation Combined in a Second-Order Scheme. J. comp. phys., 1974(14)

[72] Virand H. Formes Conservatives Equations da la Dynamiquc das Gas La Recherche Aenspatinle. 1974(1)

[73] Schlichting H. Boundary Layer Theory. New York: McGraw-Hill, 1979

[74] 高智.流体力学基本方程组的层次结构理论和简化NS方程组.力学学报,1988(2)

[75] 王正华.超音速飞行器底部喷流强干扰流场的并行数值计算[工学博士学位论文].长沙:国防科学技术大学,1992

[76] 杨晓辉.二维/三维超声速飞行器底部复杂流场的并行数值模拟研究[工学博士学位论文].长沙:国防科技大学,1997

[77] 王正华,王承尧.轴对称横向喷流强干扰流场的巨型机向量并行计算.宇航学报,1995,16(1)

[78] 王正华,王承尧.显式差分程序的巨型机宏任务并行计算.空气动力学学报,1995,13(2)

[79] Wang Zhenghua, etal. Vectorization and Parallel Computation of a CFD Code on YH-2 Parallel Supercomputer. Wuhan University Journal of Natural Sciences,1996,1(3/4)

[80] Yang Xiaohui, Wang Zhenghua, etal. Parallel Computation of Supersonic Blunt Body Viscous Flow-

fields in PVM. Chinese Journal of Aeronautics,1997,10(4)
[81] 王承尧,王正华,陈宏等. 高超声速与超声速喷流的数值计算和实验研究. 中国科学基金,1997,11(1)
[82] 杨晓辉,王承尧,王正华. PVM 环境下激波的嵌入式自适应求解. 湖南大学学报,1997,24(5)
[83] 陈军,王正华,李晓梅. 高阶稠密矩阵的并行求逆与应用. 国防科技大学学报,1997,19(4)
[84] 曾明,王正华等. 再入飞行器化学非平衡流动的隐式向量算法研究. 推进技术,1997,18(5)
[85] 李桦,王承尧,王正华等. 高超音速流场的隐式 TVD 格式并行计算. 国防科技大学学报,1998,20(2)
[86] 陈军,王正华,李晓梅. 并行计算中性能评估方法的研究. 计算机工程,1998,24(12)
[87] 杨晓辉,王承尧,王正华. 旋成体超声速底压特性的并行数值研究. 空气动力学学报,1999,17(1)
[88] 杨晓辉,王正华,王承尧. PVM 平台下通量守恒的嵌入自适应网格解激波和底部流. 空气动力学学报,1999,17(3)
[89] 赵文涛,王正华等. 并行处理方法在液体火箭发动机三维数值模拟中的应用. 国防科技大学学报,1999,21(4)
[90] 迟利华,李晓梅,王正华. 块带状线性方程组的分布式并行算法. 计算机工程与科学,1999,21(3)
[91] 杨晓辉,王正华等. PVM 平台下三维超声速钝头体底部流场的并行数值模拟及其拓扑分析. 国防科技大学学报,1998,20(6)
[92] 赵文涛,王正华等. 火箭发动机两相喷雾燃烧的并行虚拟机仿真. 推进技术,1999,20(4)
[93] 李桦,王承尧,王正华. PVM 环境下提高并行计算加速比的数值实验研究. 宇航学报,1999,20(1)
[94] 赵文涛,王正华等. 液体火箭发动机燃烧室内复杂流动模拟的并行 SIMPLE 算法. 推进技术,1999,20(6)
[95] 迟利华. 大型稀疏线性方程组在分布式存储环境下的计算[工学博士学位论文]. 长沙:国防科学技术大学,1998
[96] 王正华,车永刚,赵文涛. 三对角方程组的分布式 SPP 算法. 国防科技大学学报,1999,21(5)
[97] 贺国宏. 三阶 ENN 格式及其在高超声速粘性复杂流场求解中的应用[工学博士学位论文]. 中国空气动力研究与发展中心,1994
[98] Amodio P,Mastronardi N. A Parallel Version of the Cyclic Reduction Algorithm on a Hypercube. Parallel Computing,1993,19:1273—1281
[99] Wang H H. A Parallel Method of Tridiagonal Equations. ACM Trans. on Math Software,1981,7(2)
[100] Michielse P H,Van Vorst H A. Data Transport in Wang's Partition Method. Parallel Computing,1988,7
[101] 李晓梅,任兵,宋君强. 并行计算与偏微分方程数值解. 长沙:国防科技大学出版社,1990
[102] 张保琳,袁国兴,刘兴平,陈劲. 偏微分方程并行有限差分方法. 北京:科学出版社,1994
[103] 齐治昌. 数值分析及其应用. 长沙:国防科技大学出版社,1987
[104] Stone H. An Efficient Parallel Algorithm for the Solution of a Tridiagonal Linear System of Equations. J. ACM,1973,20
[105] Michael J Quinn. The PRAM Model of Parallel Computation. Parallel theory and practice,1994,27—30
[106] Susanne E Hambrush,Ashfag A Khokhar. C3:A Parallel Model for Coarse-Rrained Machine. J. of parallel and distributed systems,1996,32
[107] Sun X,Rover D T. Scalability of Parallel Algorithm—Machine Combinations. IEEE trans. on par-

allel and distributed systems,1994,5(6)

[108] Gustfson J, Montry G, Benner R. Development of Parallel Methods for a 1024-processor Hypercube. SIAM J. SSTC,1988(9)

[109] Kumar V,Singh V. Scalability of Parallel Algorithms for the Allpairs Shortest Path Problem:A Summary of Results. Proc. of Conf. on Parallel Processing,Chicago,IL,1990:136—140

[110] Heath M T,Etheride J A. Visualizing the Performance of Parallel Programs. IEEE Software, 1991,8